

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 5, 2002, 10:34:58 ; Search time 81.22 Seconds
(without alignments)
3053.354 Million cell updates/sec

Title: US-09-680-514-8

Perfect score: 1095

Sequence: 1 ATGGAGCTGACTGAATTGCT.....TACGCCACTTCCCGCCGCC 1095

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_NA:*

- 1: /cgn2.6/ptodata/2/ina/5A_COMB.seq:*
- 2: /cgn2.6/ptodata/2/ina/5B_COMB.seq:*
- 3: /cgn2.6/ptodata/2/ina/6A_COMB.seq:*
- 4: /cgn2.6/ptodata/2/ina/6B_COMB.seq:*
- 5: /cgn2.6/ptodata/2/ina/PTUS_COMB.seq:*
- 6: /cgn2.6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	518.8	47.4	1059	US-08-471-045-57	Sequence 57, Appl
2	518.8	47.4	1059	US-08-469-712A-57	Sequence 57, Appl
3	518.8	47.4	1062	US-08-347-029-3	Sequence 3, Appl
4	518.8	47.4	1062	US-08-457-254-3	Sequence 3, Appl
5	518.8	47.4	1062	US-08-484-257-18	Sequence 18, Appl
6	518.8	47.4	1062	US-08-482-212-1	Sequence 1, Appl
7	518.8	47.4	1062	US-08-999-927-1	Sequence 1, Appl
8	518.8	47.4	1062	PCT-US94-08806-18	Sequence 18, Appl
9	518.8	47.4	1062	PCT-US95-01829-1	Sequence 1, Appl
10	518.8	47.4	1062	PCT-US95-14932-3	Sequence 3, Appl
11	518.8	47.4	1062	PCT-US95-16626-3	Sequence 3, Appl
12	518.8	47.4	1342	US-08-388-779A-1	Sequence 1, Appl
13	518.8	47.4	1342	US-08-591-070A-1	Sequence 1, Appl
14	518.8	47.4	1342	US-08-413-803-24	Sequence 24, Appl
15	518.8	47.4	1342	US-08-321-488A-24	Sequence 24, Appl
16	518.8	47.4	1342	US-08-414-161B-1	Sequence 1, Appl
17	518.8	47.4	1342	US-08-927-855-1	Sequence 1, Appl
18	518.8	47.4	1342	PCT-US95-03776-24	Sequence 24, Appl
19	518.8	47.4	1342	PCT-US95-03776-26	Sequence 26, Appl
20	518.8	47.4	1795	US-08-330-517-1	Sequence 1, Appl
21	517.2	47.2	605	US-08-388-779A-3	Sequence 3, Appl
22	517.2	47.2	605	US-08-591-070A-3	Sequence 3, Appl
23	517.2	47.2	605	US-08-927-855-3	Sequence 3, Appl
24	486.8	44.5	525	US-08-434-411-1	Sequence 1, Appl
25	486.8	44.5	525	US-08-434-402-1	Sequence 1, Appl
26	486.8	44.5	525	US-08-783-288-1	Sequence 1, Appl
27	486.8	44.5	525	US-08-890-640-1	Sequence 1, Appl

28	486.8	44.5	525	6	5194592-25	Patent No. 5194592
29	486	44.4	921	3	US-08-469-318-60	Sequence 60, Appl
30	486	44.4	921	3	US-08-469-318-87	Sequence 87, Appl
31	486	44.4	921	3	US-08-468-609A-60	Sequence 60, Appl
32	486	44.4	921	3	US-08-468-609A-87	Sequence 87, Appl
33	486	44.4	921	5	PCT-US95-01185-60	Sequence 60, Appl
34	486	44.4	921	5	PCT-US95-01185-87	Sequence 87, Appl
35	482.8	44.1	921	3	US-08-469-318-84	Sequence 84, Appl
36	482.8	44.1	921	3	US-08-468-609A-84	Sequence 84, Appl
37	481.2	43.9	1017	3	PCT-US95-01185-84	Sequence 84, Appl
38	481.2	43.9	1017	3	US-08-469-318-67	Sequence 67, Appl
39	481.2	43.9	1017	3	US-08-468-609A-67	Sequence 67, Appl
40	481.2	43.9	1017	5	PCT-US95-01185-67	Sequence 67, Appl
41	478	43.7	921	3	US-08-469-318-56	Sequence 56, Appl
42	478	43.7	921	3	US-08-469-318-85	Sequence 85, Appl
43	478	43.7	921	3	US-08-468-609A-56	Sequence 56, Appl
44	478	43.7	921	3	US-08-468-609A-85	Sequence 85, Appl
45	478	43.7	921	5	PCT-US95-01185-56	Sequence 56, Appl

ALIGNMENTS

RESULT 1

US-08-471-045-57

; Sequence 57, Application US/08471045

; Patent No. 6060047

; GENERAL INFORMATION:

; APPLICANT: Abrams, Mark A.

; APPLICANT: Bauer, S. C.

; APPLICANT: Braford-Goldberg, Sarah R.

; APPLICANT: Caparon, Mair H.

; APPLICANT: Easton, Alan M.

; APPLICANT: Klein, Barbara K.

; APPLICANT: McKearn, John P.

; APPLICANT: Ollins, Peter O.

; APPLICANT: Paik, Kumnan

; APPLICANT: Thomas, John W.

; TITLE OF INVENTION: Co-administration of Interleukin-3

; TITLE OF INVENTION: Mutant Polypeptides with CSF's for Multi-lineage

; TITLE OF INVENTION: Hematopoietic Cell Production

; NUMBER OF SEQUENCES: 58

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Dennis A. Bennett, G.D. Searle & Co.,

; ADDRESS: Corporate Patent Dept.

; STREET: P. O. Box 5110

; CITY: Chicago

; STATE: Illinois

; COUNTRY: USA

; ZIP: 60680

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/471,045

; FILING DATE: 06-JUN-1995

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/US95/01184

; FILING DATE: 02-FEB-1995

; APPLICATION NUMBER: US 08/193,373

; FILING DATE: 04-FEB-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Bennett, Dennis A.

; REGISTRATION NUMBER: 34,547

; REFERENCE/DOCKET NUMBER: C-2789/3

; TELEPHONE: (708)470-6501

; TELEFAX: (708)470-6881

; INFORMATION FOR SEQ ID NO: 57:

SEQUENCE CHARACTERISTICS:
LENGTH: 1059 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-471-045-57

Query Match 47.4%; Score 518.8; DB 3; Length 1059;
Best Local Similarity 99.6%; Pred. No. 1.4e-116;
Matches 520; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGGAGCTGACTGAATTCCTCGTGGTCACTGCTTCTCTAACTCAAGGCTAAACGCTG 60
Db 1 ATGGAGCTGACTGAATTCCTCGTGGTCACTGCTTCTCTAACTCAAGGCTAAACGCTG 60

Qy 61 TCCAGCCCGGCTCTCTGCTTGTGACCTCCGAGTCTCAGTAACTGCTTCGTGACTCC 120
Db 61 TCCAGCCCGGCTCTCTGCTTGTGACCTCCGAGTCTCAGTAACTGCTTCGTGACTCC 120

Qy 121 CATGCTCTTCCACAGCAGCTGAGCCAGTCCAGAGGTTTCAACCTTTCCTACACCTGTC 180
Db 121 CATGCTCTTCCACAGCAGCTGAGCCAGTCCAGAGGTTTCAACCTTTCCTACACCTGTC 180

Qy 181 CTGCTGCTGCTGTGGAGCAGTACCTTCTGGAACCCAGATGGAGGAGACCAAG 240
Db 181 CTGCTGCTGCTGTGGAGCAGTACCTTCTGGAACCCAGATGGAGGAGACCAAG 240

Qy 241 GCACAGACATTTCTGGAGCAGTACCTTCTGCTGGAGGAGTATGTCGACGACGGGA 300
Db 241 GCACAGACATTTCTGGAGCAGTACCTTCTGCTGGAGGAGTATGTCGACGACGGGA 300

Qy 301 CAACCTGAGATCCCAATGCTTCTCATCCCTCTGGGAGCAGTTCCTGGACAGGTCCTGTC 360
Db 301 CAACCTGAGATCCCAATGCTTCTCATCCCTCTGGGAGCAGTTCCTGGACAGGTCCTGTC 360

Qy 481 CGTTCTCTGATGCTGTAGGAGGTCACACCTCTGGGTACGG 522
Db 481 CGTTCTCTGATGCTGTAGGAGGTCACACCTCTGGGTACGG 522

RESULT 2

US-08-469-712A-57
Sequence 57, Application US/08469712A
Patent No. 6093395
GENERAL INFORMATION:
APPLICANT: Abrams, Mark A.
APPLICANT: Bauer, S. C.
APPLICANT: Braford-Goldberg, Sarah R.
APPLICANT: Caparon, Maïre H.
APPLICANT: Easton, Alan M.
APPLICANT: Klein, Barbara K.
APPLICANT: McKeown, John P.
APPLICANT: Olin, Peter O.
APPLICANT: Paik, Kumnan
APPLICANT: Thomas, John W.
TITLE OF INVENTION: Co-administration of Interleukin-3
TITLE OF INVENTION: Mutant Polypeptides with CSF's for Multi-lineage
TITLE OF INVENTION: Hematopoietic Cell Production
NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dennis A. Bennett, G.D. Searle & Co.,
ADDRESSEE: Corporate Patent Dept.
STREET: P. O. Box 5110

CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60680
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469,712A
FILING DATE: 06-JUN-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/01184
FILING DATE: 02-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/193,373
FILING DATE: 04-FEB-1994
ATTORNEY/AGENT INFORMATION:
NAME: Bennett, Dennis A.
REGISTRATION NUMBER: 34,547
REFERENCE/DOCKET NUMBER: C-2789/2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (708)470-6501
TELEFAX: (708)470-6881
INFORMATION FOR SEQ ID NO: 57:
SEQUENCE CHARACTERISTICS:
LENGTH: 1059 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-469-712A-57

Query Match 47.4%; Score 518.8; DB 3; Length 1059;

Best Local Similarity 99.6%; Pred. No. 1.4e-116;

Matches 520; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGGAGCTGACTGAATTCCTCGTGGTCACTGCTTCTCTAACTCAAGGCTAAACGCTG 60
Db 1 ATGGAGCTGACTGAATTCCTCGTGGTCACTGCTTCTCTAACTCAAGGCTAAACGCTG 60

Qy 61 TCCAGCCCGGCTCTCTGCTTGTGACCTCCGAGTCTCAGTAACTGCTTCGTGACTCC 120
Db 61 TCCAGCCCGGCTCTCTGCTTGTGACCTCCGAGTCTCAGTAACTGCTTCGTGACTCC 120

Qy 121 CATGCTCTTCCACAGCAGCTGAGCCAGTCCAGAGGTTTCAACCTTTCCTACACCTGTC 180
Db 121 CATGCTCTTCCACAGCAGCTGAGCCAGTCCAGAGGTTTCAACCTTTCCTACACCTGTC 180

Qy 181 CTGCTGCTGCTGTGGAGCAGTACCTTCTGGAACCCAGATGGAGGAGACCAAG 240
Db 181 CTGCTGCTGCTGTGGAGCAGTACCTTCTGGAACCCAGATGGAGGAGACCAAG 240

Qy 241 GCACAGACATTTCTGGAGCAGTACCTTCTGCTGGAGGAGTATGTCGACGACGGGA 300
Db 241 GCACAGACATTTCTGGAGCAGTACCTTCTGCTGGAGGAGTATGTCGACGACGGGA 300

Qy 301 CAACCTGAGATCCCAATGCTTCTCATCCCTCTGGGAGCAGTTCCTGGACAGGTCCTGTC 360
Db 301 CAACCTGAGATCCCAATGCTTCTCATCCCTCTGGGAGCAGTTCCTGGACAGGTCCTGTC 360

Qy 361 CTCTTGGGGCCCTGCAGAGCCTTCTTGAACCCAGCTTCTCCACAGGCGAGGACCA 420
Db 361 CTCTTGGGGCCCTGCAGAGCCTTCTTGAACCCAGCTTCTCCACAGGCGAGGACCA 420

Qy 421 GCTCAAGGATCCCAATGCTTCTTCTGAGCTTCCACACCTGCTCCGAGGAAAGGTG 480
Db 421 GCTCAAGGATCCCAATGCTTCTTCTGAGCTTCCACACCTGCTCCGAGGAAAGGTG 480

Qy 481 CGTTCTCTGATGCTGTAGGAGGTCACACCTCTGCGTAGGG 522

Db 481 CGTTTCCTGATGCTTAGGAGGGTCCACCCTCTGGTCAGG 522
|||||

RESULT 3

US-08-347-029-3

; Sequence 3, Application US/08347029

; Patent No. 5641655

; GENERAL INFORMATION:

; APPLICANT: Foster, Donald C.

; APPLICANT: Heipel, Mark

; APPLICANT: Holly, Richard D.

; TITLE OF INVENTION: METHOD FOR SECRETING THROMBOPOIETIN

; TITLE OF INVENTION: POLYPEPTIDES

; NUMBER OF SEQUENCES: 15

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: ZymoGenetics, Inc.

; STREET: 1201 Eastlake Avenue East

; CITY: Seattle

; STATE: WA

; COUNTRY: USA

; ZIP: 98102

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/347,029

; FILING DATE:

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Parker, Gary E

; REGISTRATION NUMBER: 31-648

; REFERENCE/DOCKET NUMBER: 94-13

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 206-442-6673

; TELEFAX: 206-442-6678

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1062 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 1..1059

US-08-347-029-3

Query Match 47.4%; Score 518.8; DB 1; Length 1062;
Best Local Similarity 99.6%; Pred. No. 1.4e-116;
Matches 520; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGAGCTGACTGAATTCGCTCTCGTGGTCAATGCTTCTCTAAGTCAAGGCTAAGCGTG 60
|||||
Db 1 ATGGAGCTGACTGAATTCGCTCTCGTGGTCAATGCTTCTCTAAGTCAAGGCTAAGCGTG 60
|||||
QY 61 TCCAGCCCGGCTCCCTCTGCTTGTGACCTCCGAGTCCAGTAACTGCTTCGTGACTCC 120
|||||
Db 61 TCCAGCCCGGCTCCCTCTGCTTGTGACCTCCGAGTCCAGTAACTGCTTCGTGACTCC 120
|||||
QY 121 CATGTCCTTCACAGACTGAGCAGTCCAGTCCAGAGGTTACCCCTTTCCTACACTGTC 180
|||||
Db 121 CATGTCCTTCACAGACTGAGCAGTCCAGTCCAGAGGTTACCCCTTTCCTACACTGTC 180
|||||
QY 181 CTGCTGCTCTGCTGCTGCTTGTGACCTCCGAGTCCAGTAACTGCTTCGTGACTCC 240
|||||
Db 181 CTGCTGCTCTGCTGCTGCTTGTGACCTCCGAGTCCAGTAACTGCTTCGTGACTCC 240
|||||
QY 241 GCACAGGACATCTTGGGAGCAGTACCCCTTCTGCTGGAGGAGTGTATGCGACGCGGGA 300
|||||

Db 241 GCACAGGACATCTTGGGAGCAGTACCCCTTCTGCTGGAGGAGTGTATGCGACGCGGGA 300
QY 301 CAACTGGGACCCACTTGCCTCTCATCCCTCCCTGGGSCAGCTTCTCGACAGGTCCCGTCTC 360
|||||
Db 301 CAACTGGGACCCACTTGCCTCTCATCCCTCCCTGGGSCAGCTTCTCGACAGGTCCCGTCTC 360
|||||
QY 361 CTCCTTGGGGCCCTGCAGAGCCCTCTTGGAAACCCAGCTTCTCTCCACAGGCGAGGACCACA 420
|||||
Db 361 CTCCTTGGGGCCCTGCAGAGCCCTCTTGGAAACCCAGCTTCTCTCCACAGGCGAGGACCACA 420
|||||
QY 421 GCTCACAAGGATCCCAATGCCATCTTCTCGAGCTTCCACACACTGCTCCGAGGAGAAAGGTG 480
|||||
Db 421 GCTCACAAGGATCCCAATGCCATCTTCTCGAGCTTCCACACACTGCTCCGAGGAGAAAGGTG 480
|||||
QY 481 CGTTTCCTCATGCTGTAGGAGGTCACCCCTCTGCGTACGG 522
|||||
Db 481 CGTTTCCTCATGCTGTAGGAGGTCACCCCTCTGCGTACGG 522
|||||

RESULT 4

US-08-457-254-3

; Sequence 3, Application US/08457254

; Patent No. 5986049

; GENERAL INFORMATION:

; APPLICANT: Forstrom, John W

; APPLICANT: Lofton-Day, Catherine E

; APPLICANT: Lok, Si

; TITLE OF INVENTION: PURIFIED THROMBOPOIETIN AND METHOD OF

; TITLE OF INVENTION: MAKING IT

; NUMBER OF SEQUENCES: 24

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: ZymoGenetics, Inc.

; STREET: 1201 Eastlake Avenue East

; CITY: Seattle

; STATE: WA

; COUNTRY: USA

; ZIP: 98102

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/457,254

; FILING DATE:

; CLASSIFICATION: 514

; ATTORNEY/AGENT INFORMATION:

; NAME: Parker, Gary E

; REGISTRATION NUMBER: 31,648

; REFERENCE/DOCKET NUMBER: 94-11C1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 206-442-6673

; TELEFAX: 206-442-6678

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1062 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 1..1059

US-08-457-254-3

Query Match 47.4%; Score 518.8; DB 2; Length 1062;
Best Local Similarity 99.6%; Pred. No. 1.4e-116;
Matches 520; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGAGCTGACTGAATTCGCTCTCGTGGTCAATGCTTCTCTAAGTCAAGGCTAAGCGTG 60
|||||
Db 1 ATGGAGCTGACTGAATTCGCTCTCGTGGTCAATGCTTCTCTAAGTCAAGGCTAAGCGTG 60
|||||

Query Match	47.4%	Score 518.8	DB 3	Length 1062
Best Local Similarity	99.6%	Pred. No. 1.4e-116		
Matches 520	Conservative 0	Mismatches 2	Indels 0	Gaps 0
QY	1	ATGGAGCTGACTGAATTGCTCCTCGTGGTCAATGCTTCTCTAACTGCAAGGCTAACGCTG	60	
DB	1	ATGGAGCTGACTGAATTGCTCCTCGTGGTCAATGCTTCTCTAACTGCAAGGCTAACGCTG	60	
QY	61	TCCAGCCGGCTCCTCCTGCTGTGTGACCTCCGAGTCCCTCAGTAAACTGCTTCGTGACTCC	120	
DB	61	TCCAGCCGGCTCCTCCTGCTGTGTGACCTCCGAGTCCCTCAGTAAACTGCTTCGTGACTCC	120	
QY	121	CATGTCTTTACAGCAGACTGAGCCAGTGCCAGAGGTTACCCCTTTGGCTTACACCTGTC	180	
DB	121	CATGTCTTTACAGCAGACTGAGCCAGTGCCAGAGGTTACCCCTTTGGCTTACACCTGTC	180	
QY	181	CTGCTGCCTGCTGTGGACTTTAGCTTGGGAGAAATGGAACCCAGATGGAGAGACCAAG	240	
DB	181	CTGCTGCCTGCTGTGGACTTTAGCTTGGGAGAAATGGAACCCAGATGGAGAGACCAAG	240	
QY	241	GCACAGACATTCCTGGAGCAGTGACCCCTTCTGCTGGAGGGAGTGTGGCAGCAGGGGA	300	
DB	241	GCACAGACATTCCTGGAGCAGTGACCCCTTCTGCTGGAGGGAGTGTGGCAGCAGGGGA	300	
QY	301	CAACTGGAGCCACTTGCCTTCTCATCCCTCTGCTGGGCGAGCTTTCCTGGACAGGTCCTGTC	360	
DB	301	CAACTGGAGCCACTTGCCTTCTCATCCCTCTGCTGGGCGAGCTTTCCTGGACAGGTCCTGTC	360	
QY	361	CTCCTTTGGGGCCCTGCAGAGCCTCCTTTGGAACCCAGCTTCCTCCACAGGGCAGGACACA	420	
DB	361	CTCCTTTGGGGCCCTGCAGAGCCTCCTTTGGAACCCAGCTTCCTCCACAGGGCAGGACACA	420	

LOCATION: 1..1059
PCT-US95-01829-1

Query Match 47.4%; Score 518.8; DB 5; Length 1062;
Best Local Similarity 99.6%; Pred. No. 1.4e-116;
Matches 520; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAGCTGACTGAATTCCTCGTGGTCAATCTCTCTAACTCAAGGCTAAGCGTG 60
DB 1 ATGAGCTGACTGAATTCCTCGTGGTCAATCTCTCTAACTCAAGGCTAAGCGTG 60
QY 61 TCCAGCCCGGCTCCTCTGCTGTGAGCTCCGAGTCTCAGTAACTGCTTCGTGACTCC 120
DB 61 TCCAGCCCGGCTCCTCTGCTGTGAGCTCCGAGTCTCAGTAACTGCTTCGTGACTCC 120
QY 121 CATGCTCTTACAGCAGACTGAGCCAGTGCAGAGTTTACACCTTTGCTTACACCTGTC 180
DB 121 CATGCTCTTACAGCAGACTGAGCCAGTGCAGAGTTTACACCTTTGCTTACACCTGTC 180
QY 181 CTGCTGCTGCTGTGACTTTAGCTTGGGAGATGAAACCCAGATGGAGGAGACCAAG 240
DB 181 CTGCTGCTGCTGTGACTTTAGCTTGGGAGATGAAACCCAGATGGAGGAGACCAAG 240
QY 241 GCACAGGACATTCCTGAGCAGTGAACCTTCTGCTGGAGGAGTATGGCAGCAGCGGGA 300
DB 241 GCACAGGACATTCCTGAGCAGTGAACCTTCTGCTGGAGGAGTATGGCAGCAGCGGGA 300
QY 301 CAAGTGGGACCTTGCCTCTCATCCCTCTGCTGGGAGGAGTTTCTGGACAGGTCCTGTC 360
DB 301 CAAGTGGGACCTTGCCTCTCATCCCTCTGCTGGGAGGAGTTTCTGGACAGGTCCTGTC 360
QY 361 CTGCTGGGACCTTGCCTCTCATCCCTCTGCTGGGAGGAGTTTCTGGACAGGTCCTGTC 420
DB 361 CTGCTGGGACCTTGCCTCTCATCCCTCTGCTGGGAGGAGTTTCTGGACAGGTCCTGTC 420
QY 421 GCTCACAAGGATCCCAATGCCATCTTCTGAGCTTCCACACCTGCTCCGAGGAGGAGTG 480
DB 421 GCTCACAAGGATCCCAATGCCATCTTCTGAGCTTCCACACCTGCTCCGAGGAGGAGTG 480
QY 481 CGTTTCTGATGCTTAGAGGAGGTCACCCCTCTGCGTACGG 522
DB 481 CGTTTCTGATGCTTAGAGGAGGTCACCCCTCTGCGTACGG 522

RESULT 10

PCT-US95-14932-3
; Sequence 3, Application PC/TUS9514932
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; APPLICANT: 1201 Eastlake Avenue East
; APPLICANT: Seattle
; APPLICANT: WA
; APPLICANT: USA
; APPLICANT: 98102
; TITLE OF INVENTION: METHOD FOR SECRETING THROMBOPOIETIN
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/14932
; FILING DATE:

CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, Gary E
; REGISTRATION NUMBER: 31-648
; REFERENCE/DOCKET NUMBER: 94-13PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6673
; TELEFAX: 206-442-6678
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1062 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1059
; PCT-US95-14932-3

Query Match 47.4%; Score 518.8; DB 5; Length 1062;
Best Local Similarity 99.6%; Pred. No. 1.4e-116;
Matches 520; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAGCTGACTGAATTCCTCGTGGTCAATCTCTCTAACTCAAGGCTAAGCGTG 60
DB 1 ATGAGCTGACTGAATTCCTCGTGGTCAATCTCTCTAACTCAAGGCTAAGCGTG 60
QY 61 TCCAGCCCGGCTCCTCTGCTGTGAGCTCCGAGTCTCAGTAACTGCTTCGTGACTCC 120
DB 61 TCCAGCCCGGCTCCTCTGCTGTGAGCTCCGAGTCTCAGTAACTGCTTCGTGACTCC 120
QY 121 CATGCTCTTACAGCAGACTGAGCCAGTGCAGAGTTTACACCTTTGCTTACACCTGTC 180
DB 121 CATGCTCTTACAGCAGACTGAGCCAGTGCAGAGTTTACACCTTTGCTTACACCTGTC 180
QY 181 CTGCTGCTGCTGTGACTTTAGCTTGGGAGATGAAACCCAGATGGAGGAGACCAAG 240
DB 181 CTGCTGCTGCTGTGACTTTAGCTTGGGAGATGAAACCCAGATGGAGGAGACCAAG 240
QY 241 GCACAGGACATTCCTGAGCAGTGAACCTTCTGCTGGAGGAGTATGGCAGCAGCGGGA 300
DB 241 GCACAGGACATTCCTGAGCAGTGAACCTTCTGCTGGAGGAGTATGGCAGCAGCGGGA 300
QY 301 CAAGTGGGACCTTGCCTCTCATCCCTCTGCTGGGAGGAGTTTCTGGACAGGTCCTGTC 360
DB 301 CAAGTGGGACCTTGCCTCTCATCCCTCTGCTGGGAGGAGTTTCTGGACAGGTCCTGTC 360
QY 361 CTGCTGGGACCTTGCCTCTCATCCCTCTGCTGGGAGGAGTTTCTGGACAGGTCCTGTC 420
DB 361 CTGCTGGGACCTTGCCTCTCATCCCTCTGCTGGGAGGAGTTTCTGGACAGGTCCTGTC 420
QY 421 GCTCACAAGGATCCCAATGCCATCTTCTGAGCTTCCACACCTGCTCCGAGGAGGAGTG 480
DB 421 GCTCACAAGGATCCCAATGCCATCTTCTGAGCTTCCACACCTGCTCCGAGGAGGAGTG 480
QY 481 CGTTTCTGATGCTTAGAGGAGGTCACCCCTCTGCGTACGG 522
DB 481 CGTTTCTGATGCTTAGAGGAGGTCACCCCTCTGCGTACGG 522

RESULT 11

PCT-US95-16626-3
; Sequence 3, Application PC/TUS9516626
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; APPLICANT: 1201 Eastlake Avenue East
; APPLICANT: Seattle
; APPLICANT: WA
; APPLICANT: USA
; APPLICANT: 98102
; TITLE OF INVENTION: PURIFIED THROMBOPOIETIN AND METHOD OF

;; TITLE OF INVENTION: MAKING IT
;; NUMBER OF SEQUENCES: 24
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: ZymoGenetics, Inc.
;; STREET: 1201 Eastlake Avenue East
;; CITY: Seattle
;; STATE: WA
;; COUNTRY: USA
;; ZIP: 98102
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: PCT/US95/16626
;; FILING DATE:
;; CLASSIFICATION:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Parker, Gary E
;; REGISTRATION NUMBER: 31,648
;; REFERENCE/DOCKET NUMBER: 94-11PC
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 206-442-6673
;; TELEFAX: 206-442-6678
;; INFORMATION FOR SEQ ID NO: 3:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1062 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; MOLECULE TYPE: cdna
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: 1..1059
;; PCT-US95-16626-3

Query Match 47.4%; Score 518.8; DB 5; Length 1062;
Best Local Similarity 99.6%; Pred. No. 1.4e-116;
Matches 520; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGAGCTGACTGAATTGCTCTCGTGGTCAATGCTTCTCTAACTGCAAGGCTAACGCTG 60
Db 1 ATGAGCTGACTGAATTGCTCTCGTGGTCAATGCTTCTCTAACTGCAAGGCTAACGCTG 60
Qy 61 TCCAGCCCGGCTCTCTCTGCTGACCTCCGAGTCCCTCAGTAACTGCTTCGTGACTCC 120
Db 61 TCCAGCCCGGCTCTCTCTGCTGACCTCCGAGTCCCTCAGTAACTGCTTCGTGACTCC 120
Qy 121 CATGTCTCTCACAGCAGCTGAGCCAGTGCAGGAGTTTCAACCTTTGCTTACACTGTC 180
Db 121 CATGTCTCTCACAGCAGCTGAGCCAGTGCAGGAGTTTCAACCTTTGCTTACACTGTC 180
Qy 181 CTGCTGCTCTGTGAGCTTTAGCTTTGGAGATGAAACCCAGATGGAGGAGACCAAG 240
Db 181 CTGCTGCTCTGTGAGCTTTAGCTTTGGAGATGAAACCCAGATGGAGGAGACCAAG 240
Qy 241 GCACAGACATCTGGAGCAGTGAACCTTCTGCTGGAGGAGTGTATGCGACAGGGGA 300
Db 241 GCACAGACATCTGGAGCAGTGAACCTTCTGCTGGAGGAGTGTATGCGACAGGGGA 300
Qy 301 CAACCTGGAGCCCACTTGCCTCTCATCCCTCTCTGGGAGCTTTCTGAGACAGTCCGCTC 360
Db 301 CAACCTGGAGCCCACTTGCCTCTCATCCCTCTCTGGGAGCTTTCTGAGACAGTCCGCTC 360
Qy 361 CTGCTTTGGGCCCCGAGAGCCCTTCTTGAACCCAGCTTCTTCCACAGGGCAGGACCA 420
Db 361 CTGCTTTGGGCCCCGAGAGCCCTTCTTGAACCCAGCTTCTTCCACAGGGCAGGACCA 420
Qy 421 GCTCACAAGGATCCCAATGCCATCTTCTGAGCTTCCCAACCTGCTCCGAGGAAAGGTG 480
Db 421 GCTCACAAGGATCCCAATGCCATCTTCTGAGCTTCCCAACCTGCTCCGAGGAAAGGTG 480

Qy 481 CGTTTCTCTGATGCTCTAGGAGGTTCCACCTCTCGGTACGG 522
Db 481 CGTTTCTCTGATGCTCTAGGAGGTTCCACCTCTCGGTACGG 522
RESULT 12
US-08-388-779A-1
; Sequence 1, Application US/08388779A
; Patent No. 5696250
; GENERAL INFORMATION:
; APPLICANT: Elliott, Steven G.
; TITLE OF INVENTION: MGF ANALOGS
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: AMGEN INC.
; STREET: 1840 Behavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320-1720
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/388,779A
; FILING DATE: 15-FEB-1995
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Cook Ph.D., Robert R.
; REGISTRATION NUMBER: 31,602
; REFERENCE/DOCKET NUMBER: A-337
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1342 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 36..1094
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 99..1094
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 36..98
; US-08-388-779A-1

Query Match 47.4%; Score 518.8; DB 1; Length 1342;
Best Local Similarity 99.6%; Pred. No. 1.5e-116;
Matches 520; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGAGCTGACTGAATTGCTCTCGTGGTCAATGCTTCTCTAACTGCAAGGCTAACGCTG 60
Db 36 ATGAGCTGACTGAATTGCTCTCGTGGTCAATGCTTCTCTAACTGCAAGGCTAACGCTG 95
Qy 61 TCCAGCCCGGCTCTCTCTGCTGACCTCCGAGTCCCTCAGTAACTGCTTCGTGACTCC 120
Db 96 TCCAGCCCGGCTCTCTCTGCTGACCTCCGAGTCCCTCAGTAACTGCTTCGTGACTCC 155
Qy 121 CATGTCTCTCACAGCAGCTGAGCCAGTGCAGGAGTTTCAACCTTTGCTTACACTGTC 180
Db 156 CATGTCTCTCACAGCAGCTGAGCCAGTGCAGGAGTTTCAACCTTTGCTTACACTGTC 215
Qy 181 CTGCTGCTCTGTGAGCTTTAGCTTTGGAGATGAAACCCAGATGGAGGAGACCAAG 240
Db 216 CTGCTGCTCTGTGAGCTTTAGCTTTGGAGATGAAACCCAGATGGAGGAGACCAAG 275

1	ATGAGCTGACTGAATTGCTCTCGTGGTCACTGCTTCTCACTGCAAGGCTAACGCTG	60
36	ATGAGCTGACTGAATTGCTCTCGTGGTCACTGCTTCTCACTGCAAGGCTAACGCTG	95
61	TCCAGCCCGCTCCTCTGCTGTGACCTCCGAGTCCCTCAGTAAACTGCTTCGTGACTCC	120
96	TCCAGCCCGCTCCTCTGCTGTGACCTCCGAGTCCCTCAGTAAACTGCTTCGTGACTCC	155
121	CATGTCCTTACACGACAGCTGAGCCAGTGCCTCAGAGGTTCACTTCCTTACACCTGTC	180
156	CATGTCCTTACACGACAGCTGAGCCAGTGCCTCAGAGGTTCACTTCCTTACACCTGTC	215
181	CTGCTGCCCTGCTGAGCACTTTAGCTTGGGAGAAATGGAAACCCAGATGGAGGACCAAG	240
216	CTGCTGCCCTGCTGAGCACTTTAGCTTGGGAGAAATGGAAACCCAGATGGAGGACCAAG	275
241	GCACAGGACATTCCTGGGACAGTAGCCCTTCTGCTGGAGGAGTATGGCAGCACGGGA	300
276	GCACAGGACATTCCTGGGACAGTAGCCCTTCTGCTGGAGGAGTATGGCAGCACGGGA	335
301	CAACTGGGACCCACTTGCCTCTCATCCCTCTCGGGGACAGTTTCTTGACAGGTCCTCTC	360
336	CAACTGGGACCCACTTGCCTCTCATCCCTCTCGGGGACAGTTTCTTGACAGGTCCTCTC	395
361	CTCCTTTGGGCCCCTGCAGAGCCTCCTTTGGAACCCAGTTCCTCCACAGGCGAGGCCACA	420
396	CTCCTTTGGGCCCCTGCAGAGCCTCCTTTGGAACCCAGTTCCTCCACAGGCGAGGCCACA	455
421	GCTCACAAAGGATCCCAATGCCATCTTCCTGAGCTTCCACACCTGCTCCGAGGAAGGTG	480
456	GCTCACAAAGGATCCCAATGCCATCTTCCTGAGCTTCCACACCTGCTCCGAGGAAGGTG	515
481	CGTTTCTCTGATGCTTTAGAGAGGTCACCTCTCGGTACGG	522
516	CGTTTCTCTGATGCTTTAGAGAGGTCACCTCTCGGTACGG	557

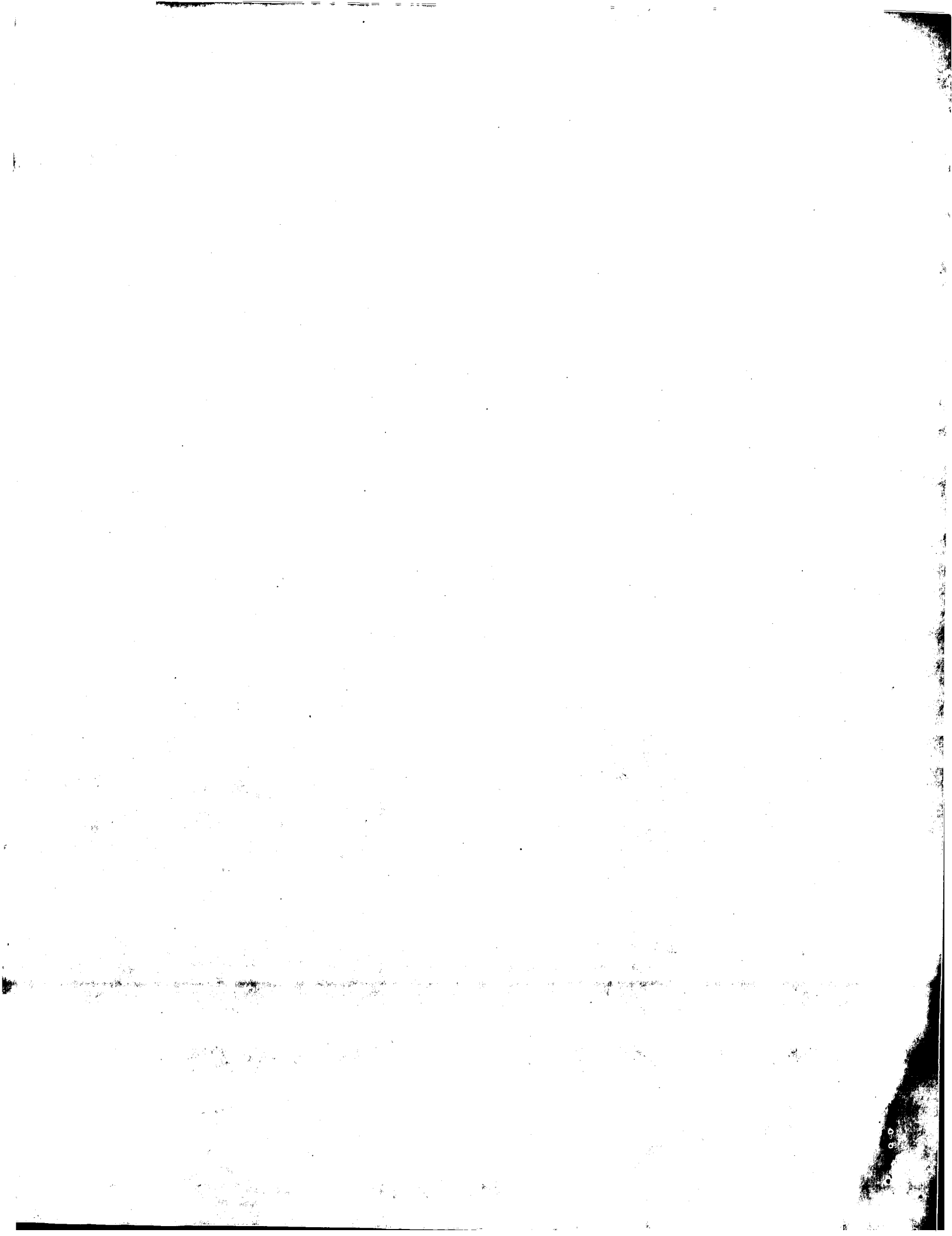
Query Match	47.4%;	Score 518.8;	DB 1;	Length 1342;
Best Local Similarity	99.6%;	Pred. No. 1.5e-116;		
Matches 520;	Conservative	0;	Mismatches 2;	Indels 0;
Gaps	0;			
QY	1	ATGGAGCTGACTGAATTGCTTCCTCGTGGTCATCGTCTCTTAACATGCAAGSCTAACCGCTG	60	
Db	36	ATGGAGCTGACTGAATTGCTTCCTCGTGGTCATCGTCTCTTAACATGCAAGSCTAACCGCTG	95	
QY	61	TCAGCCCGGGTCCTCCTCGTGTGACCTCCGAGTCCTCAGTAAACTGCTTCGTGACTCC	120	
Db	96	TCAGCCCGGGTCCTCCTCGTGTGACCTCCGAGTCCTCAGTAAACTGCTTCGTGACTCC	155	
QY	121	CATGTCCCTTCACAGCAGACTGACCCAGTCGCCAGAGGTTTCACCCTTGGCTACACCGTGC	180	

```

Db 156 CATGTCCTTCACAGCAGACTGAGCCAGTGCCAGAGGTTACCCCTTTGCTACACCTGTC 215
Qy 181 CTGCTGCTGCTGTGGACTTTAGCTTGGAGAAATGGAACCCACAGATGGAGGAGACCAAG 240
Db 216 CTGCTGCTGCTGTGGACTTTAGCTTGGAGAAATGGAACCCACAGATGGAGGAGACCAAG 275
Qy 241 GCACAGGACATTCCTGGGAGCAGTGACCCCTTCTGTGGAGGAGTGATGGCAGCAGCGGGA 300
Db 276 GCACAGGACATTCCTGGGAGCAGTGACCCCTTCTGTGGAGGAGTGATGGCAGCAGCGGGA 335
Qy 301 CAACTGGGACCCACTTGGCTCTCATCCCTCTGCTGGGCGAGCTTCTGGGACAGGTCGCTC 360
Db 336 CAACTGGGACCCACTTGGCTCTCATCCCTCTGCTGGGCGAGCTTCTGGGACAGGTCGCTC 395
Qy 361 CTCTTGGGGCCCTGCAGAGCTCTCTGGAACCCAGCTTCTCCACAGGCGAGGACCACA 420
Db 396 CTCTTGGGGCCCTGCAGAGCTCTCTGGAACCCAGCTTCTCCACAGGCGAGGACCACA 455
Qy 421 GCTCACAAGGATCCCAATGCCATCTTCTGAGCTTCCAAACACCTGCTCCGAGGAAAGGTG 480
Db 456 GCTCACAAGGATCCCAATGCCATCTTCTGAGCTTCCAAACACCTGCTCCGAGGAAAGGTG 515
Qy 481 CGTTTCCTGATGCTGTAGGAGGTTCCACCCTCTCGGTACGG 522
Db 516 CGTTTCCTGATGCTGTAGGAGGTTCCACCCTCTCGGTACGG 557

```

Search completed: April 5, 2002, 10:35:04
 Job time: 4998 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 5, 2002, 10:34:53 ; Search time 81.22 Seconds
(without alignments)
3019.893 Million cell updates/sec

Title: US-09-680-514-6

Perfect score: 1083

Sequence: 1 ATGGAGCTGACTGAATTGCT.....TACGCCACCTGCCAGCC 1083

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_NA.*
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCTUS.COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	523.2	48.3	1059	3	US-08-471-045-57
2	523.2	48.3	1059	3	US-08-469-712A-57
3	523.2	48.3	1062	1	US-08-347-029-3
4	523.2	48.3	1062	2	US-08-457-254-3
5	523.2	48.3	1062	2	US-08-484-257-18
6	523.2	48.3	1062	3	US-08-482-212-1
7	523.2	48.3	1062	3	US-08-999-927-1
8	523.2	48.3	1062	5	PCT-US94-08806-18
9	523.2	48.3	1062	5	PCT-US95-01829-1
10	523.2	48.3	1062	5	PCT-US95-14932-3
11	523.2	48.3	1062	5	PCT-US95-16626-3
12	523.2	48.3	1342	1	US-08-388-779A-1
13	523.2	48.3	1342	1	US-08-591-070A-1
14	523.2	48.3	1342	1	US-08-413-803-24
15	523.2	48.3	1342	1	US-08-321-488A-24
16	523.2	48.3	1342	2	US-08-414-161B-1
17	523.2	48.3	1342	2	US-08-927-855-1
18	523.2	48.3	1342	5	PCT-US95-03776-24
19	523.2	48.3	1342	5	PCT-US95-03776-26
20	523.2	48.3	1795	1	US-08-330-517-1
21	521.6	48.2	605	1	US-08-388-779A-3
22	521.6	48.2	605	1	US-08-591-070A-3
23	521.6	48.2	605	2	US-08-927-855-3
24	486.8	44.9	525	1	US-08-434-411-1
25	486.8	44.9	525	1	US-08-434-402-1
26	486.8	44.9	525	1	US-08-783-288-1
27	486.8	44.9	525	2	US-08-890-640-1

28	486.8	44.9	525	6	5194592-25	Patent No. 5194592
29	486.6	44.9	1017	3	US-08-469-318-67	Sequence 67, Appl
30	486.6	44.9	1017	3	US-08-468-609A-67	Sequence 67, Appl
31	486.6	44.9	1017	5	PCT-US95-01185-67	Sequence 67, Appl
32	478	44.1	1164	1	US-08-413-803-26	Sequence 26, Appl
33	478	44.1	1164	1	US-08-321-488A-26	Sequence 28, Appl
34	478	44.1	1164	5	PCT-US95-03776-28	Sequence 28, Appl
35	477.6	44.1	921	3	US-08-469-318-84	Sequence 84, Appl
36	477.6	44.1	921	3	US-08-468-609A-84	Sequence 84, Appl
37	477.6	44.1	921	5	PCT-US95-01185-84	Sequence 84, Appl
38	476.2	44.0	546	3	US-08-489-318-177	Sequence 177, App
39	476.2	44.0	546	3	US-08-469-318-178	Sequence 178, App
40	476.2	44.0	546	3	US-08-468-609A-177	Sequence 177, App
41	476.2	44.0	546	4	US-08-468-609A-178	Sequence 178, App
42	476.2	44.0	546	5	PCT-US95-01185-177	Sequence 177, App
43	476.2	44.0	546	5	PCT-US95-01185-178	Sequence 178, App
44	476.2	44.0	921	3	US-08-489-318-72	Sequence 72, Appl
45	476.2	44.0	921	3	US-08-489-318-75	Sequence 75, Appl

ALIGNMENTS

RESULT 1

US-08-471-045-57
; Sequence 57, Application US/08471045
; Patent No. 6060047
; GENERAL INFORMATION:
; APPLICANT: Abrams, Mark A.
; APPLICANT: Bauer, S. C.
; APPLICANT: Braford-Goldberg, Sarah R.
; APPLICANT: Caparon, Mairé H.
; APPLICANT: Easton, Alan M.
; APPLICANT: Klein, Barbara K.
; APPLICANT: McKearn, John P.
; APPLICANT: Olin, Peter O.
; APPLICANT: Paik, Kumnan
; APPLICANT: Thomas, John W.
; TITLE OF INVENTION: Co-administration of Interleukin-3
; TITLE OF INVENTION: Mutant Polypeptides with CSP's for Multi-lineage
; TITLE OF INVENTION: Hematopoietic Cell Production
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dennis A. Bennett, G.D. Searle & Co.,
; ADDRESSEE: Corporate Patent Dept.
; STREET: P. O. Box 5110
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60680
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/471,045
; FILING DATE: 06-JUN-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/01184
; FILING DATE: 02-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/193,373
; FILING DATE: 04-FEB-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Bennett, Dennis A.
; REGISTRATION NUMBER: 34,547
; REFERENCE/DOCKET NUMBER: C-2789/3
; TELEPHONE: (708)470-6501
; TELEFAX: (708)470-6881
; INFORMATION FOR SEQ ID NO: 57:

Q:

```

: TITLE OF INVENTION: Using Hematopoietic Proteins.
:
: NUMBER OF SEQUENCES: 6
:
: CORRESPONDENCE ADDRESS:
:
: ADDRESSEE: zymoGenetics, Inc.
:
: STREET: 1201 Eastlake Avenue East
:
: CITY: Seattle
:
: STATE: WA
:
: COUNTRY: USA
:
: ZIP: 98102
:
: COMPUTER READABLE FORM:
:
: MEDIUM TYPE: Floppy disk
:
: COMPUTER: IBM PC compatible
:
: OPERATING SYSTEM: PC-DOS/MS-DOS
:
: SOFTWARE: Patent In Release #1.0, Version #1.25
:
: CURRENT APPLICATION DATA:
:
: APPLICATION NUMBER: US/08/999,927
:
: FILING DATE:
:
: CLASSIFICATION:
:
: PRIOR APPLICATION DATA:
:
: APPLICATION NUMBER: US/08/347,748
:
: FILING DATE:
:
: ATTORNEY/AGENT INFORMATION:
:
: NAME: Parker, Garv. P.
:

```

REFERENCE/DOCKET NUMBER: 94-0962
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-442-6673
TELEFAX: 206-442-6678

```

; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1062 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1059
; US-08-999-927-1
;
Query Match          48.3%; Score 523.2; DB 3; Length 1062;
Best Local Similarity 99.4%; Pred. No. 31e-118;
Matches 525; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ATGAGCTGACTGAATTCCTCGTGGTCTATGCTTCTCCTAACTGCAAGCTTAACGCTG 60
    |||
Db 1 ATGAGCTGACTGAATTCCTCGTGGTCTATGCTTCTCCTAACTGCAAGCTTAACGCTG 60
    |||

Qy 61 TCCAGCCCGGCTCCTCCTGCTTGAGACCTCGGAGTCTCTCAGTAAACTCGTTTCGFGACTCC 120
    |||
Db 61 TCCAGCCCGGCTCCTCCTGCTTGAGACTCTCGAGTCTCTCAGTAAACTCGTTTCGFGACTCC 120
    |||

Qy 121 CATGTCCTTCACAGCAGACTGAGCCAGTGCCTGAGAGGTTTCAACCTTTTGCCTTACACCTGTG 180
    |||
Db 121 CATGTCCTTCACAGCAGACTGAGCCAGTGCCTGAGAGGTTTCAACCTTTTGCCTTACACCTGTG 180
    |||

Qy 181 CTGCTGCCTCTGTGGAGCTTACGTTGGGAGAAATGGAACCCAGATGGAGGAGACCAAG 240
    |||
Db 181 CTGCTGCCTCTGTGGAGCTTACGTTGGGAGAAATGGAACCCAGATGGAGGAGACCAAG 240
    |||

Qy 241 GCACAGGACATTCGAGGAGCAGTGACCTTCTGCTGGAGGAGTGATGGCAGCACGGGGA 300
    |||
Db 241 GCACAGGACATTCGAGGAGCAGTGACCTTCTGCTGGAGGAGTGATGGCAGCACGGGGA 300
    |||

```

Query Match	48.3%	Score 523.2	DB 3	Length 1062
Best Local Similarity	99.4%	Pred. No. 3.1e-118		
Matches 525	Conservative 0	Mismatches 3	Indels 0	Gaps 0
QY 1	ATGGAGCTGACTGAATTCCTCGTGGTGCATGCTTCTCCTAACTGCAAGGCTAAACGCTG	60		
DB 1	ATGGAGCTGACTGAATTCCTCGTGGTGCATGCTTCTCCTAACTGCAAGGCTAAACGCTG	60		
QY 61	TCCAGCCCGGCTCCTCTGTTGTGACCTCCGAGTCCCTCAGTAAACTGCTTCGTGACTCC	120		
DB 61	TCCAGCCCGGCTCCTCTGTTGTGACCTCCGAGTCCCTCAGTAAACTGCTTCGTGACTCC	120		
QY 121	CATGTCCTTCACAGCAGACTGACCCAGTGCCTCAGAGAGTTACCCCTTTTGCTTACACCTGTC	180		
DB 121	CATGTCCTTCACAGCAGACTGACCCAGTGCCTCAGAGAGTTACCCCTTTTGCTTACACCTGTC	180		
QY 181	CTGCTGCCTCTGTGGACTTTAGCTTGGGAGATGAAAAACCCAGATGGAGGACCAAG	240		
DB 181	CTGCTGCCTCTGTGGACTTTAGCTTGGGAGATGAAAAACCCAGATGGAGGACCAAG	240		
QY 241	GCACAGACATTTGCGGAGCAGTGACCCCTTCTGCTGGAGGGAGTGATGGCAGCAGCGGGA	300		
DB 241	GCACAGACATTTGCGGAGCAGTGACCCCTTCTGCTGGAGGGAGTGATGGCAGCAGCGGGA	300		
QY 301	CAACTGGGACCCACTTGCCTCTCATCCCTCCTGGGGCAGCTTCTTGACAGAGTCCGCTCTC	360		
DB 301	CAACTGGGACCCACTTGCCTCTCATCCCTCCTGGGGCAGCTTCTTGACAGAGTCCGCTCTC	360		
QY 361	CTCCTTTGGGGCCCTGACAGAGCTCCTTTGGAACCCAGCTTCTCCTCACAGGGCAGGACACA	420		
DB 361	CTCCTTTGGGGCCCTGACAGAGCTCCTTTGGAACCCAGCTTCTCCTCACAGGGCAGGACACA	420		

```
OY 421 GCTCACAAGGATCCCAATGCTTCCTGAGCTTCCAAACACCTGCTCCGAGGAAAGGTG 480
Db 421 GCTCACAAGGATCCCAATGCTTCCTGAGCTTCCAAACACCTGCTCCGAGGAAAGGTG 480
OY 481 CGTTTCTGATGCTGTAGGAGGTCCACCTCTGCGTCAGGGGTGCG 528
Db 481 CGTTTCTGATGCTGTAGGAGGTCCACCTCTGCGTCAGGGGTGCG 528

RESULT 8
PCT-US94-08806-18
; Sequence 18, Application PC/TUS9408806
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; APPLICANT: 1201 Eastlake Avenue East
; APPLICANT: Seattle
; APPLICANT: WA
; APPLICANT: US
; APPLICANT: 98102
; APPLICANT: APPLICANT: University of Washington
; APPLICANT: Seattle
; APPLICANT: WA
; APPLICANT: 98195
; TITLE OF INVENTION: HEMATOPOIETIC PROTEIN AND MATERIALS AND
; TITLE OF INVENTION: METHODS FOR MAKING IT
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/08806
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, Gary E
; REGISTRATION NUMBER: 31-648
; REFERENCE/DOCKET NUMBER: 93-12PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6600 ext 6673
; TELEFAX: 206-442-6678
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1062 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1059
PCT-US94-08806-18

Query Match 48.3%; Score 523.2; DB 5; Length 1062;
Best Local Similarity 99.4%; Pred. No. 3.1e-118;
Matches 525; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 ATGAGCTGACTGAATGCTCCTCGGTGCTATGCTTCTTAACCTCAAGGCTAACGCTG 60
Db 1 ATGAGCTGACTGAATGCTCCTCGGTGCTATGCTTCTTAACCTCAAGGCTAACGCTG 60
OY 61 TCCAGCCCGGCTCCTCCTGCTGTGACCTCCGAGTCCCTCAGTAACTGCTTGTGACTCC 120
Db 61 TCCAGCCCGGCTCCTCCTGCTGTGACCTCCGAGTCCCTCAGTAACTGCTTGTGACTCC 120
```

```
OY 121 CATGTCTCTTACAGCAGACTGAGCCAGTCCAGAGGTTTCCACCTTTGCTTACACCTGTC 180
Db 121 CATGTCTCTTACAGCAGACTGAGCCAGTCCAGAGGTTTCCACCTTTGCTTACACCTGTC 180
OY 181 CTGCTGCTGCTGTGAGACTTTAGCTTTGGGAGAAATGAAACCCAGATGGAGGAGACCAAG 240
Db 181 CTGCTGCTGCTGTGAGACTTTAGCTTTGGGAGAAATGAAACCCAGATGGAGGAGACCAAG 240
OY 241 GCACAGGACATTTCTGGGAGCAGTGACCTTCTCTGCTGAGGAGTGTATGGCAGACGGGA 300
Db 241 GCACAGGACATTTCTGGGAGCAGTGACCTTCTCTGCTGAGGAGTGTATGGCAGACGGGA 300
OY 301 CAACTGGGACCCACTTGCCTCTCATCCCTCTCTGGGAGAGCTTTCTTGACAGAGTCCGCTC 360
Db 301 CAACTGGGACCCACTTGCCTCTCATCCCTCTCTGGGAGAGCTTTCTTGACAGAGTCCGCTC 360
OY 361 CTCTCTGGGGCCCTGAGAGAGCTTCTTGGAAACCCAGCTTCTCCACAGGCGAGGACCA 420
Db 361 CTCTCTGGGGCCCTGAGAGAGCTTCTTGGAAACCCAGCTTCTCCACAGGCGAGGACCA 420
OY 421 GCTCAAGGATCCCAATGCCATCTTCTCAGCTTCCAAACACCTGCTCCGAGGAAAGGTG 480
Db 421 GCTCAAGGATCCCAATGCCATCTTCTCAGCTTCCAAACACCTGCTCCGAGGAAAGGTG 480
OY 481 CGTTTCTGATGCTGTAGGAGGTCCACCTCTGCTGCTCAGGGGTGCG 528
Db 481 CGTTTCTGATGCTGTAGGAGGTCCACCTCTGCTGCTCAGGGGTGCG 528

RESULT 9
PCT-US95-01829-1
; Sequence 1, Application PC/TUS9501829
; GENERAL INFORMATION:
; APPLICANT: University of Washington
; APPLICANT: Seattle
; APPLICANT: WA
; APPLICANT: 98195
; TITLE OF INVENTION: Methods of Stimulating Erythropoiesis
; TITLE OF INVENTION: Using Hematopoietic Proteins.
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/01829
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, Gary E
; REGISTRATION NUMBER: 31-648
; REFERENCE/DOCKET NUMBER: 94-09PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6673
; TELEFAX: 206-442-6678
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1062 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
```

```

; LOCATION: 1..1059
PCT-US95-01829-1

Query Match      48.3%; Score 523.2; DB 5; Length 1062;
Best Local Similarity 99.4%; Pred. No. 3.1e-118;
Matches 525; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGAGCTGACTGAATGCTCTCGTGTGCTATGCTTCTCCTAACTGCAAGGCTAACGCTG 60
Db 1 ATGAGCTGACTGAATGCTCTCGTGTGCTATGCTTCTCCTAACTGCAAGGCTAACGCTG 60
QY 61 TCCAGCCGGCTCTCTGCTGTGACCTCGAGTCCCTCAGTAACTGCTTCGTGACTCC 120
Db 61 TCCAGCCGGCTCTCTGCTGTGACCTCGAGTCCCTCAGTAACTGCTTCGTGACTCC 120
QY 121 CATGCTCTTACAGCAGACTGAGCGAGTGGCCAGAGGTTTACCCCTTGCCTTACACCTGTC 180
Db 121 CATGCTCTTACAGCAGACTGAGCGAGTGGCCAGAGGTTTACCCCTTGCCTTACACCTGTC 180
QY 181 CTGCTGCTGCTGTGGACTTTAGCTTGGGAGAAATGGAAACCCAGATGGAGGAGACCAAG 240
Db 181 CTGCTGCTGCTGTGGACTTTAGCTTGGGAGAAATGGAAACCCAGATGGAGGAGACCAAG 240
QY 241 GCACAGGACATTTCTGGAGCAGTGAACCTTCTGCTGGAGGAGTGAATGGCAGCAGCGGGA 300
Db 241 GCACAGGACATTTCTGGAGCAGTGAACCTTCTGCTGGAGGAGTGAATGGCAGCAGCGGGA 300
QY 301 CAACCTGGACCCACTTGGCTCTCATCCCTCTGCGGGCAGCTTCTGACAGGTCGCTGTC 360
Db 301 CAACCTGGACCCACTTGGCTCTCATCCCTCTGCGGGCAGCTTCTGACAGGTCGCTGTC 360
QY 361 CTCCTTGGGGCCCTGACAGAGCTCTTGGAAACCCAGCTTCTCCACAGGGCAGGACCA 420
Db 361 CTCCTTGGGGCCCTGACAGAGCTCTTGGAAACCCAGCTTCTCCACAGGGCAGGACCA 420
QY 421 GCTCACAAGGATCCCAATGCCATCTTCTGAGCTTCCAACACCTGCTCCGAGGAAAGGTG 480
Db 421 GCTCACAAGGATCCCAATGCCATCTTCTGAGCTTCCAACACCTGCTCCGAGGAAAGGTG 480
QY 481 CGTTTCTGATGCTTGTAGGAGGTCACCCCTCTGCGTCAAGGGTGGC 528
Db 481 CGTTTCTGATGCTTGTAGGAGGTCACCCCTCTGCGTCAAGGGTGGC 528

RESULT 11
PCT-US95-14932-3
; Sequence 3, Application PC/TUS9514932
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; APPLICANT: 1201 Eastlake Avenue East
; APPLICANT: Seattle
; APPLICANT: WA
; APPLICANT: USA
; APPLICANT: 98102
; TITLE OF INVENTION: METHOD FOR SECRETING THROMBOPOIETIN
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/14932
; FILING DATE:
; TITLE OF INVENTION: PURIFIED THROMBOPOIETIN AND METHOD OF
```

```

; TITLE OF INVENTION: MAKING IT
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: .USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/16626
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, Gary E
; REGISTRATION NUMBER: 31,648
; REFERENCE/DOCKET NUMBER: 94-11PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6673
; TELEFAX: 206-442-6678
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1062 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1059
; PCT-US95-16626-3

```

```

Query Match      48.3%; Score 523.2; DB 5; Length 1062;
Best Local Similarity 99.4%; Pred. No. 3.1e-118;
Matches 525; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 1 ATGGAGCTGACTGAATTGCTCCTCGTGTGATGCTTCTCTAACTCAAGGCTAACGCTG 60
Db 1 ATGGAGCTGACTGAATTGCTCCTCGTGTGATGCTTCTCTAACTCAAGGCTAACGCTG 60
QY 61 TCCAGCCCGGCTCCTCGTGTGATGCTTCTCTAACTCAAGGCTAACGCTG 120
Db 61 TCCAGCCCGGCTCCTCGTGTGATGCTTCTCTAACTCAAGGCTAACGCTG 120
QY 121 CATGTCCTTTCACAGCAGCTGAGCCAGTCCAGAGGTTTCAACCTTTCCTACACCTGTC 180
Db 121 CATGTCCTTTCACAGCAGCTGAGCCAGTCCAGAGGTTTCAACCTTTCCTACACCTGTC 180
QY 181 CTGTCGCTGTGTGGACTTTAGCTTGGGAGATGGAAACCCAGATGGAGGAGACCAAG 240
Db 181 CTGTCGCTGTGTGGACTTTAGCTTGGGAGATGGAAACCCAGATGGAGGAGACCAAG 240
QY 241 GCACAGACATCTTGGGAGCAGTGACCTTCTCTCGGAGGAGTGTATGCGACACGGGA 300
Db 241 GCACAGACATCTTGGGAGCAGTGACCTTCTCTCGGAGGAGTGTATGCGACACGGGA 300
QY 301 CAACCTGGGACCCACTTGCCTTCTCATCCCTCCTGGGAGCTTTCCTGGAGAGTCCGCTC 360
Db 301 CAACCTGGGACCCACTTGCCTTCTCATCCCTCCTGGGAGCTTTCCTGGAGAGTCCGCTC 360
QY 361 CTCTTGGGCGCTCGAGAGCCTTCTTGAACCCAGCTTCTTCCACAGGAGGAGACCA 420
Db 361 CTCTTGGGCGCTCGAGAGCCTTCTTGAACCCAGCTTCTTCCACAGGAGGAGACCA 420
QY 421 GCTCACAAGGATCCCAATGCCATCTTCTGAGCTTCCAAACCTCTGCTCCGAGGAAGTG 480
Db 421 GCTCACAAGGATCCCAATGCCATCTTCTGAGCTTCCAAACCTCTGCTCCGAGGAAGTG 480

```

```

QY 481 CGTTTCCTGATGCTGTAGGAGGTTCCACCTCTCGTCAAGGGTGGC 528
Db 481 CGTTTCCTGATGCTGTAGGAGGTTCCACCTCTCGTCAAGGGTGGC 528

```

RESULT 12

```

US-08-388-779A-1
; Sequence 1, Application US/08388779A
; Patent No. 5696250
; GENERAL INFORMATION:
; APPLICANT: Elliott, Steven G.
; TITLE OF INVENTION: MGDF ANALOGS
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: AMGEN INC.
; STREET: 1840 Dehavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320-1720
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/388,779A
; FILING DATE: 15-FEB-1995
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Cook Ph.D., Robert R.
; REGISTRATION NUMBER: 31,602
; REFERENCE/DOCKET NUMBER: A-337
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1342 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 36..1094
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 99..1094
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 36..98
; US-08-388-779A-1

```

```

Query Match      48.3%; Score 523.2; DB 1; Length 1342;
Best Local Similarity 99.4%; Pred. No. 3.3e-118;
Matches 525; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 1 ATGGAGCTGACTGAATTGCTCCTCGTGTGATGCTTCTCTAACTCAAGGCTAACGCTG 60
Db 36 ATGGAGCTGACTGAATTGCTCCTCGTGTGATGCTTCTCTAACTCAAGGCTAACGCTG 95
QY 61 TCCAGCCCGGCTCCTCGTGTGATGCTTCTCTAACTCAAGGCTAACGCTG 120
Db 96 TCCAGCCCGGCTCCTCGTGTGATGCTTCTCTAACTCAAGGCTAACGCTG 155
QY 121 CATGTCCTTTCACAGCAGCTGAGCCAGTCCAGAGGTTTCAACCTTTCCTACACCTGTC 180
Db 156 CATGTCCTTTCACAGCAGCTGAGCCAGTCCAGAGGTTTCAACCTTTCCTACACCTGTC 215
QY 181 CTGTCGCTGTGTGGACTTTAGCTTGGGAGATGGAAACCCAGATGGAGGAGACCAAG 240
Db 216 CTGTCGCTGTGTGGACTTTAGCTTGGGAGATGGAAACCCAGATGGAGGAGACCAAG 275

```


QY 241 GCACAGGACATTCGGGAGCAGTACACCTTCTGCTGGAGGAGTGATGGCAGCAGCGGGA 300
Db 276 GCACAGGACATTCGGGAGCAGTACACCTTCTGCTGGAGGAGTGATGGCAGCAGCGGGA 335
QY 301 CAACCTGGGACCCACTTGGCTCTCATCCCTCTGCTGGGAGAGCTTCTGGACAGGTCCGTCTC 360
Db 336 CAACCTGGGACCCACTTGGCTCTCATCCCTCTGCTGGGAGAGCTTCTGGACAGGTCCGTCTC 395
QY 361 CTCCTTGGGGCCCTGCAGAGCCTCTTGGAAACCCAGCTTCTCCACAGGGGAGGACACACA 420
Db 396 CTCCTTGGGGCCCTGCAGAGCCTCTTGGAAACCCAGCTTCTCCACAGGGGAGGACACACA 455
QY 421 GCTCACAGGATCCCAATGCCATCTTCTGAGCTTCCAAACACCTGCTCCGAGGAAAGGTG 480
Db 456 GCTCACAGGATCCCAATGCCATCTTCTGAGCTTCCAAACACCTGCTCCGAGGAAAGGTG 515
QY 481 CGTTTCTGTAGCTTGTAGGAGGCTCCACCTCTGCGTCAGGGGTGGC 528
Db 516 CGTTTCTGTAGCTTGTAGGAGGCTCCACCTCTGCGTCAGGGGTGGC 563

RESULT 13

US-08-591-070A-1
; Sequence 1, Application US/08591070A
; Patent No. 5756083
; GENERAL INFORMATION:
; APPLICANT: Elliott, Steven G.
; TITLE OF INVENTION: MPL LIGAND ANALOGS
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: AMGEN INC.
; STREET: 1840 Dehavilland Drive
; CITY: Thousand Oaks
; STATE: CA
; COUNTRY: USA
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/591,070A
; FILING DATE: 09-FEB-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/388,779
; FILING DATE: 15-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: COOK, ROBERT R
; REGISTRATION NUMBER: 31,602
; REFERENCE/DOCKET NUMBER: A-337B
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1342 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 36..1094
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 99..1094
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 36..98
US-08-591-070A-1

Query Match 48.3%; Score 523.2; DB 1; Length 1342;
Best Local Similarity 99.4%; Pred. No. 3.3e-118;

Matches 525; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 ATGAGCTGACTGAATTCGCTCCTCGTGTATGCTTCTCCTAACTGCAAGGCTAACCGTG 60
Db 36 ATGAGCTGACTGAATTCGCTCCTCGTGTATGCTTCTCCTAACTGCAAGGCTAACCGTG 95
QY 61 TCCAGCCGGCTCCTCCTGCTTGTGACCTCCGAGTCCCTCAGTAAACTGCTTCGTGACTCC 120
Db 96 TCCAGCCGGCTCCTCCTGCTTGTGACCTCCGAGTCCCTCAGTAAACTGCTTCGTGACTCC 155
QY 121 CATGCTCTTACACAGACTGAGCCAGTGGCCAGAGGTTCACCCCTTGGCTTACACCTGTCT 180
Db 156 CATGCTCTTACACAGACTGAGCCAGTGGCCAGAGGTTCACCCCTTGGCTTACACCTGTCT 215
QY 181 CTGCTGCTGCTGTGGAGCTTGTAGCTTGGGAAATGGAAACCCAGATGGAGGAGACCAAG 240
Db 216 CTGCTGCTGCTGTGGAGCTTGTAGCTTGGGAAATGGAAACCCAGATGGAGGAGACCAAG 275
QY 241 GCACAGGACATTCCTGGAGCAGTACCCCTTCTGCTGGAGGAGTGATGGCAGCAGCGGGA 300
Db 276 GCACAGGACATTCCTGGAGCAGTACCCCTTCTGCTGGAGGAGTGATGGCAGCAGCGGGA 335
QY 301 CACTGGGACCCACTTGCCTCTCATCCCTCTGCTGGGAGAGCTTCTGGACAGGTCCGTCTC 360
Db 336 CACTGGGACCCACTTGCCTCTCATCCCTCTGCTGGGAGAGCTTCTGGACAGGTCCGTCTC 395
QY 361 CTCCTTGGGGCCCTGCAGAGCCTCTTGGAAACCCAGCTTCTCCACAGGGGAGGACACACA 420
Db 396 CTCCTTGGGGCCCTGCAGAGCCTCTTGGAAACCCAGCTTCTCCACAGGGGAGGACACACA 455
QY 421 GCTCACAGGATCCCAATGCCATCTTCTGAGCTTCCAAACACCTGCTCCGAGGAAAGGTG 480
Db 456 GCTCACAGGATCCCAATGCCATCTTCTGAGCTTCCAAACACCTGCTCCGAGGAAAGGTG 515
QY 481 CGTTTCTGTAGCTTGTAGGAGGCTCCACCTCTGCGTCAGGGGTGGC 528
Db 516 CGTTTCTGTAGCTTGTAGGAGGCTCCACCTCTGCGTCAGGGGTGGC 563

RESULT 14

US-08-413-803-24
; Sequence 24, Application US/08413803
; Patent No. 5766581
; GENERAL INFORMATION:
; APPLICANT: Bartley, Timothy D.
; APPLICANT: Bogenberger, Jakob M.
; APPLICANT: Bosselman, Robert A.
; APPLICANT: Hunt, Pamela
; APPLICANT: Kinstler, Olaf B.
; APPLICANT: Samal, Babru B.
; TITLE OF INVENTION: METHODS FOR TREATING MAMMALS WITH
; TITLE OF INVENTION: MONO-PGYLATED PROTEINS THAT STIMULATE MEGAKARYOCYTE
; TITLE OF INVENTION: GROWTH AND DIFFERENTIATION
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: AMGEN INC.
; STREET: 1840 Dehavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: US
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/413,803
; FILING DATE: 30-MAR-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/221,768
; FILING DATE: 31-MAR-1994

```

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/252,628
FILING DATE: 31-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/321,488
FILING DATE: 12-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/347,780
FILING DATE: 30-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Cook Ph.D., Robert R.
REGISTRATION NUMBER: 31,602
REFERENCE/DOCKET NUMBER: A-290D
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 1542 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 36..1097
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 99..1097
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 36..98
US-08-413-803-24

```

Query Match	48.3%	Score 523.2	DB 1	Length 1342
Best Local Similarity	99.4%	Pred. No. 3.3e-118		
Matches 525	Conservative 0	Mismatches 0	Indels 0	Gaps 0
Qy 1	ATGGAGCTGACGTGAATTCCTCCTCGTGGTCATGCTTCTCTAACTCAAGGCTAAGCGTG	60		
Db 36	ATGGAGCTGACGTGAATTCCTCCTCGTGGTCATGCTTCTCTAACTCAAGGCTAAGCGTG	95		
Qy 61	TCCAGCCCGGCTCCTCCTGCTTGTGACCTCCGAGTCCTCAGTAAACTGCTTCGTGACTCC	120		
Db 96	TCCAGCCCGGCTCCTCCTGCTTGTGACCTCCGAGTCCTCAGTAAACTGCTTCGTGACTCC	155		
Qy 121	CATGTCCTTCACAGCAGACTGAGCCAGTGCCAGAGGTTACCCCTTTGGCTACACCTGTC	180		
Db 156	CATGTCCTTCACAGCAGACTGAGCCAGTGCCAGAGGTTACCCCTTTGGCTACACCTGTC	215		
Qy 181	CTGCTGCCCTGCTGGACTTTAGCTTTGGGAGAATGGAAACCCAGATGGAGGAGACCAAG	240		
Db 216	CTGCTGCCCTGCTGGACTTTAGCTTTGGGAGAATGGAAACCCAGATGGAGGAGACCAAG	275		
Qy 241	GCACGAGCATTTCTGGGAGCAGTGACCCCTTCTGCTGGAGGAGTGATGGCAGCACGGGA	300		
Db 276	GCACGAGCATTTCTGGGAGCAGTGACCCCTTCTGCTGGAGGAGTGATGGCAGCACGGGA	335		
Qy 301	CAACTGGGACCACATTTGCGCTCTCATCCCTCTCGGGGAGCTTTCTGGACAGGTCCGTCTC	360		
Db 336	CAACTGGGACCACATTTGCGCTCTCATCCCTCTCGGGGAGCTTTCTGGACAGGTCCGTCTC	395		
Qy 361	CTCCTTTGGGGCCCTGCAGAGCCTCTTTGGAACCCAGCTTCTCCACAGGCGAGGACCA	420		
Db 396	CTCCTTTGGGGCCCTGCAGAGCCTCTTTGGAACCCAGCTTCTCCACAGGCGAGGACCA	455		
Qy 421	GCTCAAGAAGTCCCAATGCCATCTTCTGTAGCTTCCAAACACCTGCTCCGAGGAAGGTG	480		
Db 456	GCTCAAGAAGTCCCAATGCCATCTTCTGTAGCTTCCAAACACCTGCTCCGAGGAAGGTG	515		
Qy 481	CGTTTCCCTGATGCTTGTAGGAGGGTCCACCCCTCTGCGTCAAGGGGTGGC	528		
Db 516	CGTTTCCCTGATGCTTGTAGGAGGGTCCACCCCTCTGCGTCAAGGGGTGGC	563		

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999

```

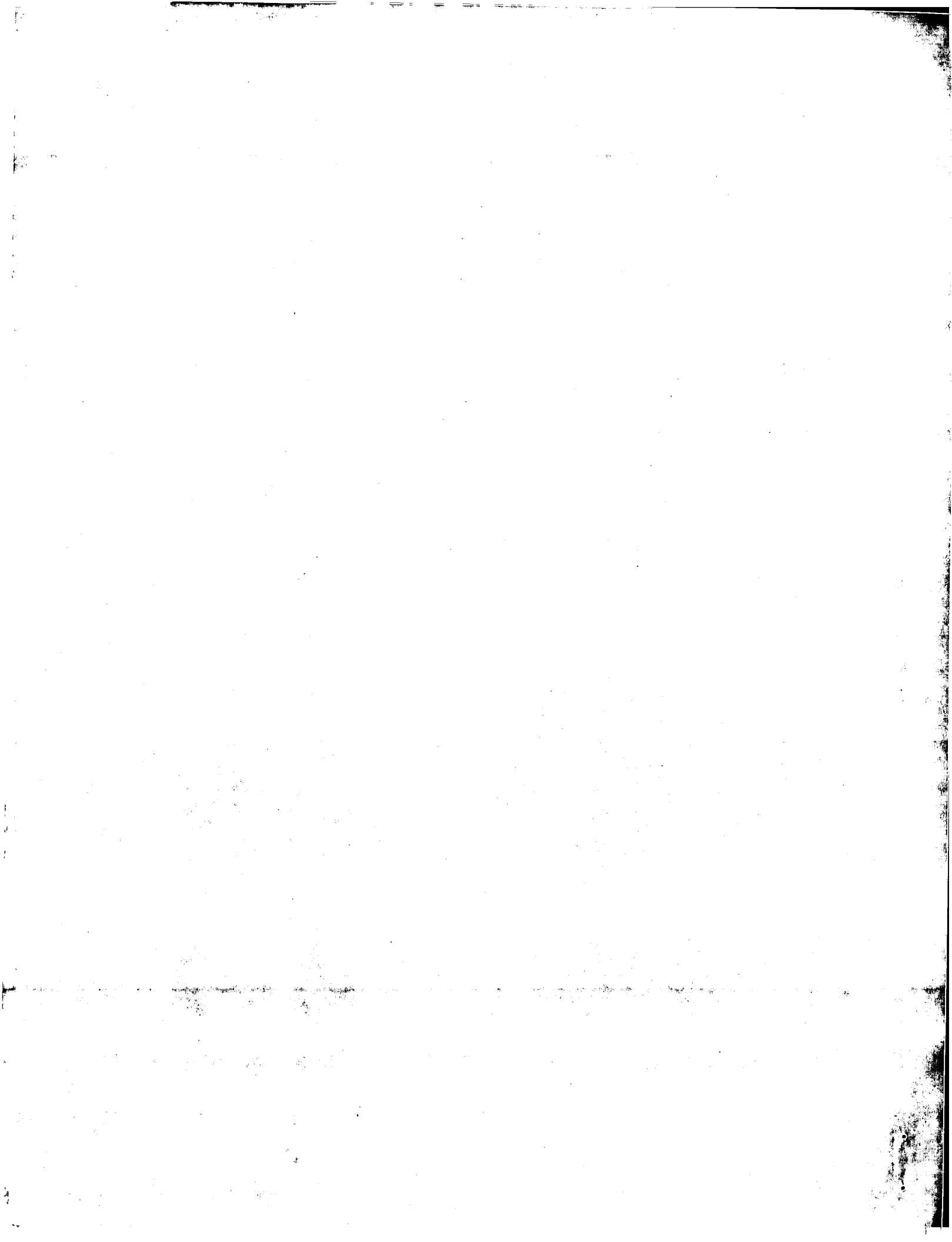
	Query Match	48.3%	Score 523.2;	DB 1;	Length 1342;
	Best Local Similarity	99.4%	Pred. No. 3.3e-118;		
	Matches 525;	Conservative	0;	Mismatches 3;	Indels 0;
	Gaps	0;			
Qy	1	ATGGAGCTGACTGAATTGCTCTCGTGGTCATGCTTCTCTAACTGCAAGGCTAACGCTG	60		
Db	36	ATGGAGCTGACTGAATTGCTCTCGTGGTCATGCTTCTCTAACTGCAAGGCTAACGCTG	95		
Qy	61	TCCAGCCGGCTCCTCTGCTTGTGACCTCGAGTCCCTCAGTAAACTGCTTCGTGACTCC	120		
Db	96	TCCAGCCGGCTCCTCTGCTTGTGACCTCGAGTCCCTCAGTAAACTGCTTCGTGACTCC	155		
Qy	121	CATGTCTCTTACAGCAGACTGAGCCAGTGCGCCAGAGGTTACCCCTTTGCCCTACACCTGTC	180		

```

Db 156 CATGTCCTTCACAGCAGACTGAGCCAGTGGCCAGAGGTTACCCCTTTCCTACACCTGTC 215
Qy 181 CTGCTGCCCTGCTGTGGACTTTAGCTTGGAGAAATGGAACCCAGATGGAGGAGACCACAG 240
Db 216 CTGCTGCCCTGCTGTGGACTTTAGCTTGGAGAAATGGAACCCAGATGGAGGAGACCACAG 275
Qy 241 GCACAGGACATTTCTGGGAGCAGTACCCTTCTGCTGGAGGAGTGATGGCAGCAGCGGGGA 300
Db 276 GCACAGGACATTTCTGGGAGCAGTACCCTTCTGCTGGAGGAGTGATGGCAGCAGCGGGGA 335
Qy 301 CAACCTGGGACCCACTTGGCTCTCATCCCTCTGGGGCAGCTTCTGGACAGGTCCGCTCTC 360
Db 336 CAACCTGGGACCCACTTGGCTCTCATCCCTCTGGGGCAGCTTCTGGACAGGTCCGCTCTC 395
Qy 361 CTCCTTGGGGCCCTGCAGAGCCTCTTGGAAACCCAGCTTCTCCACAGGGCAGGACCACA 420
Db 396 CTCCTTGGGGCCCTGCAGAGCCTCTTGGAAACCCAGCTTCTCCACAGGGCAGGACCACA 455
Qy 421 GCTCACAAAGGATCCCAATGCCATCTTCTGAGCTTCCAAACCTCTCCGAGGAAAGGTG 480
Db 456 GCTCACAAAGGATCCCAATGCCATCTTCTGAGCTTCCAAACCTCTCCGAGGAAAGGTG 515
Qy 481 CGTTTCCTGTAGCTGTAGGAGGGTCCACCCTCTGCGTCAAGGGGTGGC 528
Db 516 CGTTTCCTGTAGCTGTAGGAGGGTCCACCCTCTGCGTCAAGGGGTGGC 563

```

Search completed: April 5, 2002, 10:34:58
 Job time: 4992 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 5, 2002, 09:11:46 ; Search time 81.22 Seconds
(without alignments)
2919.508 Million cell updates/sec

Title: US-09-680-514-4

Perfect score: 1047

Sequence: 1 ATGCGAGCTGACTGAATTGCT.....TACGCCACTTGCCAGCCC 1047

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA.*

- 1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
- 2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
- 3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
- 4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
- 5: /cgn2_6/ptodata/2/ina/PTUS_COMB.seq.*
- 6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	526.6	50.3	1059	US-08-471-045-57	Sequence 57, Appl
2	526.6	50.3	1059	US-08-469-712A-57	Sequence 57, Appl
3	526.6	50.3	1062	US-08-347-029-3	Sequence 3, Appli
4	526.6	50.3	1062	US-08-457-254-3	Sequence 3, Appli
5	526.6	50.3	1062	US-08-484-257-18	Sequence 18, Appl
6	526.6	50.3	1062	US-08-482-212-1	Sequence 1, Appli
7	526.6	50.3	1062	US-08-999-927-1	Sequence 1, Appli
8	526.6	50.3	1062	PCT-US94-08806-18	Sequence 18, Appl
9	526.6	50.3	1062	PCT-US95-01829-1	Sequence 1, Appli
10	526.6	50.3	1062	PCT-US95-14932-3	Sequence 3, Appli
11	526.6	50.3	1062	PCT-US95-16626-3	Sequence 3, Appli
12	526.6	50.3	1342	US-08-388-779A-1	Sequence 1, Appli
13	526.6	50.3	1342	US-08-591-070A-1	Sequence 1, Appli
14	526.6	50.3	1342	US-08-413-803-24	Sequence 24, Appl
15	526.6	50.3	1342	US-08-321-488A-24	Sequence 24, Appl
16	526.6	50.3	1342	US-08-414-161B-1	Sequence 1, Appli
17	526.6	50.3	1342	US-08-927-855-1	Sequence 1, Appli
18	526.6	50.3	1342	PCT-US95-03776-24	Sequence 24, Appl
19	526.6	50.3	1342	PCT-US95-03776-26	Sequence 26, Appl
20	526.6	50.3	1795	US-08-330-517-1	Sequence 1, Appli
21	525	50.1	605	US-08-388-779A-3	Sequence 3, Appli
22	525	50.1	605	US-08-591-070A-3	Sequence 3, Appli
23	525	50.1	605	US-08-927-855-3	Sequence 3, Appli
24	486.8	46.5	525	US-08-434-411-1	Sequence 1, Appli
25	486.8	46.5	525	US-08-434-402-1	Sequence 1, Appli
26	486.8	46.5	525	US-08-783-288-1	Sequence 1, Appli
27	486.8	46.5	525	US-08-890-640-1	Sequence 1, Appli

28	486.8	46.5	525	6	5194592-25	Patent No. 5194592
29	478	45.7	1164	1	US-08-413-803-26	Sequence 26, Appl
30	478	45.7	1164	1	US-08-321-488A-26	Sequence 26, Appl
31	478	45.7	1164	5	PCT-US95-03776-28	Sequence 28, Appl
32	474.6	45.3	546	3	US-08-469-318-177	Sequence 177, App
33	474.6	45.3	546	3	US-08-469-318-177	Sequence 177, App
34	474.6	45.3	546	3	US-08-468-609A-177	Sequence 177, App
35	474.6	45.3	546	3	US-08-468-609A-178	Sequence 178, App
36	474.6	45.3	546	5	PCT-US95-01185-177	Sequence 177, App
37	474.6	45.3	546	5	PCT-US95-01185-178	Sequence 178, App
38	474.6	45.3	921	3	US-08-469-318-72	Sequence 72, Appl
39	474.6	45.3	921	3	US-08-469-318-75	Sequence 75, Appl
40	474.6	45.3	921	3	US-08-469-318-78	Sequence 78, Appl
41	474.6	45.3	921	3	US-08-469-318-84	Sequence 84, Appl
42	474.6	45.3	921	3	US-08-468-609A-72	Sequence 72, Appl
43	474.6	45.3	921	3	US-08-468-609A-75	Sequence 75, Appl
44	474.6	45.3	921	3	US-08-468-609A-78	Sequence 78, Appl
45	474.6	45.3	921	3	US-08-468-609A-84	Sequence 84, Appl

ALIGNMENTS

RESULT 1

US-08-471-045-57

; Sequence 57, Application US/08471045

; Patent No. 6060047

; GENERAL INFORMATION:

; APPLICANT: Abrams, Mark A.

; APPLICANT: Bauer, S. C.

; APPLICANT: Braford-Goldberg, Sarah R.

; APPLICANT: Caparon, Mair H.

; APPLICANT: Easton, Alan M.

; APPLICANT: Klein, Barbara K.

; APPLICANT: McKeown, John P.

; APPLICANT: Olin, Peter O.

; APPLICANT: Paik, Kumnan

; APPLICANT: Thomas, John W.

; TITLE OF INVENTION: Co-administration of Interleukin-3

; TITLE OF INVENTION: Mutant Polypeptides with CSF's for Multi-lineage

; TITLE OF INVENTION: Hematopoietic Cell Production

; NUMBER OF SEQUENCES: 58

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Dennis A. Bennett, G.D. Searle & Co.,

; ADDRESSEE: Corporate Patent Dept.

; STREET: P. O. Box 5110

; CITY: Chicago

; STATE: Illinois

; COUNTRY: USA

; ZIP: 60680

; COMPUTER READABLE FORM: disk

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/471,045

; FILING DATE: 06-JUN-1995

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/US95/01184

; FILING DATE: 02-FEB-1995

; PRIOR APPLICATION DATA: US 08/193,373

; APPLICATION NUMBER: 04-FEB-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Bennett, Dennis A.

; REGISTRATION NUMBER: 34,547

; REFERENCE/DOCKET NUMBER: C-2789/3

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (708)470-6501

; TELEFAX: (708)470-6881

; INFORMATION FOR SEQ ID NO: 57:

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 1059 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-471-045-57

```

```

Query Match          50.3%; Score 526.6; DB 3; Length 1059;
Best Local Similarity 99.2%; Pred. No. 2.1e-119;
Matches 529; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGGAGCTGACTGAATTCCTCGTGGTCACTCTCTTAACGCAAGGCTAACGCTG 60
Db 1 ATGGAGCTGACTGAATTCCTCGTGGTCACTCTCTTAACGCAAGGCTAACGCTG 60
QY 61 TCCAGCCCGGCTCCTCTGCTGTGACCTCCAGTCCAGTCCAGTAACTGCTTCG 120
Db 61 TCCAGCCCGGCTCCTCTGCTGTGACCTCCAGTCCAGTCCAGTAACTGCTTCG 120
QY 121 CATGCTCTTACAGCAGACTGAGCCAGTCCAGAGGTTACCCCTTTGCCCTAC 180
Db 121 CATGCTCTTACAGCAGACTGAGCCAGTCCAGAGGTTACCCCTTTGCCCTAC 180
QY 181 CTGCTGCTGCTGCTGACCTTTAGCTTGGAGAGTGGAGAAATGGAGAGCAAG 240
Db 181 CTGCTGCTGCTGCTGACCTTTAGCTTGGAGAGTGGAGAAATGGAGAGCAAG 240
QY 241 GCACAGGACATCTTGGAGCAGTACCTTCTGCTGGAGGAGTGGAGAGCAAG 300
Db 241 GCACAGGACATCTTGGAGCAGTACCTTCTGCTGGAGGAGTGGAGAGCAAG 300
QY 301 CAACTGGGACCACTTGGAGCAGTACCTTCTGCTGGAGGAGTGGAGAGCAAG 360
Db 301 CAACTGGGACCACTTGGAGCAGTACCTTCTGCTGGAGGAGTGGAGAGCAAG 360
QY 361 CTCCTTGGGGCCCTCGAGAGGCTTCCCTTGGAGGAGTGGAGAGCAAG 420
Db 361 CTCCTTGGGGCCCTCGAGAGGCTTCCCTTGGAGGAGTGGAGAGCAAG 420
QY 421 GCTCACAAGGATCCCAATGCAATCTCTCTGAGCTTCCACACCTGCTCCGAG 480
Db 421 GCTCACAAGGATCCCAATGCAATCTCTCTGAGCTTCCACACCTGCTCCGAG 480
QY 481 CGTTTCTCTGATGCTTGTAGGAGGGTCCACCTCTCGGTGAGGCGGGCCACC 533
Db 481 CGTTTCTCTGATGCTTGTAGGAGGGTCCACCTCTCGGTGAGGCGGGCCACC 533

```

```

RESULT 2
US-08-469-712A-57
; Sequence 57, Application US/08469712A
; Patent No. 6093395
; GENERAL INFORMATION:
; APPLICANT: Abrams, Mark A.
; APPLICANT: Bauer, S. C.
; APPLICANT: Bradford-Goldberg, Sarah R.
; APPLICANT: Caparon, Maïre H.
; APPLICANT: Easton, Alan M.
; APPLICANT: Klein, Barbara K.
; APPLICANT: McKearn, John P.
; APPLICANT: Ollins, Peter O.
; APPLICANT: Paik, Kumnan
; APPLICANT: Thomas, John W.
; TITLE OF INVENTION: Co-administration of Interleukin-3
; TITLE OF INVENTION: Mutant Polypeptides with CSP's for Multi-lineage
; TITLE OF INVENTION: Hematopoietic Cell Production
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dennis A. Bennett, G.D. Searle & Co.,
; ADDRESSEE: Corporate Patent Dept.
; STREET: P. O. Box 5110

```

```

; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60680
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,712A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/011184
; FILING DATE: 02-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/193,373
; FILING DATE: 04-FEB-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Bennett, Dennis A.
; REGISTRATION NUMBER: 34,547
; REFERENCE/DOCKET NUMBER: C-2789/2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (708)470-8501
; TELEFAX: (708)470-6881
; INFORMATION FOR SEQ ID NO: 57:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1059 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-469-712A-57

```

```

Query Match          50.3%; Score 526.6; DB 3; Length 1059;
Best Local Similarity 99.2%; Pred. No. 2.1e-119;
Matches 529; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGGAGCTGACTGAATTCCTCTGCTGCTGCTCTCTCTCTCTCTCTCTCTCTCT 60
Db 1 ATGGAGCTGACTGAATTCCTCTGCTGCTGCTCTCTCTCTCTCTCTCTCTCTCT 60
QY 61 TCCAGCCCGGCTCCTCTGCTGTGACCTCCAGTCCAGTCCAGTAACTGCTTCG 120
Db 61 TCCAGCCCGGCTCCTCTGCTGTGACCTCCAGTCCAGTCCAGTAACTGCTTCG 120
QY 121 CATGCTCTTACAGCAGACTGAGCCAGTCCAGAGGTTACCCCTTTGCCCTAC 180
Db 121 CATGCTCTTACAGCAGACTGAGCCAGTCCAGAGGTTACCCCTTTGCCCTAC 180
QY 181 CTGCTGCTGCTGCTGAGCTTTAGCTTGGAGAGTGGAGAAATGGAGAGCAAG 240
Db 181 CTGCTGCTGCTGCTGAGCTTTAGCTTGGAGAGTGGAGAAATGGAGAGCAAG 240
QY 241 GCACAGGACATCTTGGAGCAGTACCTTCTGCTGGAGGAGTGGAGAGCAAG 300
Db 241 GCACAGGACATCTTGGAGCAGTACCTTCTGCTGGAGGAGTGGAGAGCAAG 300
QY 301 CAACTGGGACCACTTGGAGCAGTACCTTCTGCTGGAGGAGTGGAGAGCAAG 360
Db 301 CAACTGGGACCACTTGGAGCAGTACCTTCTGCTGGAGGAGTGGAGAGCAAG 360
QY 361 CTCCTTGGGGCCCTCGAGAGGCTTCCCTTGGAGGAGTGGAGAGCAAG 420
Db 361 CTCCTTGGGGCCCTCGAGAGGCTTCCCTTGGAGGAGTGGAGAGCAAG 420
QY 421 GCTCACAAGGATCCCAATGCAATCTCTCTGAGCTTCCACACCTGCTCCGAG 480
Db 421 GCTCACAAGGATCCCAATGCAATCTCTCTGAGCTTCCACACCTGCTCCGAG 480
QY 481 CGTTTCTCTGATGCTTGTAGGAGGGTCCACCTCTCGGTGAGGCGGGCCACC 533

```

Db 481 CGTTTCTGATGCTTGTAGGAGGTCACCCCTCTGCTCAGGGGGCCCCACC 533

RESULT 3
US-08-347-029-3
; Sequence 3, Application US/08347029
; Patent No. 5641655
; GENERAL INFORMATION:
; APPLICANT: Foster, Donald C.
; APPLICANT: Heipel, Mark
; APPLICANT: Holly, Richard D.
; TITLE OF INVENTION: METHOD FOR SECRETING THROMBOPOIETIN
; TITLE OF INVENTION: POLYPEPTIDES
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: US/08/347,029
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, Gary E
; REGISTRATION NUMBER: 31-648
; REFERENCE/DOCKET NUMBER: 94-13
; TELEPHONE: 206-442-6673
; TELEFAX: 206-442-6678
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1062 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1059
US-08-347-029-3

Query Match 50.3%; Score 526.6; DB 1; Length 1062;
Best Local Similarity 99.2%; Pred. No. 2.1e-119;
Matches 529; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 ATGAGCTGACTGAATTTGCTCCGTTGGTTCATGCTTCTCTCACTCAAGGCTAACGCTG 60
Db 1 ATGAGCTGACTGAATTTGCTCCGTTGGTTCATGCTTCTCTCACTCAAGGCTAACGCTG 60
QY 61 TCAGCCCGGCTCCTCCTGCTTGTGACCTCCGAGTCCCTCAGTAACTGCTTCGTGACTCC 120
Db 61 TCAGCCCGGCTCCTCCTGCTTGTGACCTCCGAGTCCCTCAGTAACTGCTTCGTGACTCC 120
QY 121 CATGTCCTTACAGCAGACTGACCCAGTCCAGAGGTTCCACCTTTGCTACACCTGTC 180
Db 121 CATGTCCTTACAGCAGACTGACCCAGTCCAGAGGTTCCACCTTTGCTACACCTGTC 180
QY 181 CTGCTCCCTGCTGTGACTTTAGCTTGGGAGATGGAACCCAGATGGAGGAGACCAAG 240
Db 181 CTGCTCCCTGCTGTGACTTTAGCTTGGGAGATGGAACCCAGATGGAGGAGACCAAG 240
QY 241 GCACAGACATTTGAGGACAGTGACCCCTTCTGCTGGAGGAGTGTGCGACGACGCGGA 300

Db 241 GCACAGACATTTGAGGAGCAGTGACCCCTTCTGCTGAGGAGTGTGAGCAGCACGGGA 300
QY 301 CAATGGGACCCACTTTCCTCTCATCCCTCTGCGGAGCTTCTTGACAGGTCCTGCTC 360
Db 301 CAATGGGACCCACTTTCCTCTCATCCCTCTGCGGAGCTTCTTGACAGGTCCTGCTC 360
QY 361 CTCCTTGGGCGCCTGTCAGAGCCCTTGTGAACCCAGCTTCTTCCACAGGCGAGGACCA 420
Db 361 CTCCTTGGGCGCCTGTCAGAGCCCTTGTGAACCCAGCTTCTTCCACAGGCGAGGACCA 420
QY 421 GCTCACAAGGATCCCAATGCCATCTTCTGAGCTTCCACACACCTGCTCCAGGAAAGTG 480
Db 421 GCTCACAAGGATCCCAATGCCATCTTCTGAGCTTCCACACACCTGCTCCAGGAAAGTG 480
QY 481 CGTTTCTGATGCTTGTAGGAGGTCACCCCTCTGCTGAGGCGGCGGCGCAAC 533
Db 481 CGTTTCTGATGCTTGTAGGAGGTCACCCCTCTGCTGAGGCGGCGGCGGCGGCGGCGG 533

RESULT 4
US-08-457-254-3
; Sequence 3, Application US/08457254
; Patent No. 5986049
; GENERAL INFORMATION:
; APPLICANT: Forstrom, John W
; APPLICANT: Lofton-Day, Catherine E
; APPLICANT: Lok, Si
; TITLE OF INVENTION: PURIFIED THROMBOPOIETIN AND METHOD OF
; TITLE OF INVENTION: MAKING IT
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/457,254
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, Gary E
; REGISTRATION NUMBER: 31,648
; REFERENCE/DOCKET NUMBER: 94-11C1
; TELEPHONE: 206-442-6673
; TELEFAX: 206-442-6678
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1062 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1059
US-08-457-254-3

Query Match 50.3%; Score 526.6; DB 2; Length 1062;
Best Local Similarity 99.2%; Pred. No. 2.1e-119;
Matches 529; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 ATGAGCTGACTGAATTTGCTCCGTTGGTTCATGCTTCTCTCACTCAAGGCTAACGCTG 60
Db 1 ATGAGCTGACTGAATTTGCTCCGTTGGTTCATGCTTCTCTCACTCAAGGCTAACGCTG 60

```

RESULT      6
US-08-482-212-1
; Sequence 1, Application US/08482212
; Patent No. 6013067
; GENERAL INFORMATION:
; APPLICANT: Fibbe, Willem E.
; APPLICANT: Grossmann, Angelika
; TITLE OF INVENTION: METHODS FOR INCREASING HEMATOPOIETIC CELLS
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```

6
RESULT
US-08-482-
; Sequence
; Patent N°
; GENERAL
; APPLI
; APPLI
; TITLE
; NUMBER
; CORRE
; ADDR
; STRS
; CITY
; STAT
; COUN
; ZIP
; COMPUT
; MED
; COME
; OPEF

SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/482,212
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Parker, Gary E
REGISTRATION NUMBER: 31-648
REFERENCE/DOCKET NUMBER: 95-10
TELEPHONE: 206-442-6600 ext 6673
TELEFAX: 206-442-6678
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1062 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cdna
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1059
US-08-482-212-1

Query Match 50.3%; Score 526.6; DB 3; Length 1062;
Best Local Similarity 99.2%; Pred. No. 2.1e-119;
Matches 529; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGGAGCTGACTGAATGCTCTCGTGGTCAATGCTTCTCCTAACTCAAGGCTAACGCTG 60
DB 1 ATGGAGCTGACTGAATGCTCTCGTGGTCAATGCTTCTCCTAACTCAAGGCTAACGCTG 60
QY 61 TCCAGCCCGGCTCCTCTGCTTGTGACCTCCGAGTCTCAGTAACTGCTTCGTGACTCC 120
DB 61 TCCAGCCCGGCTCCTCTGCTTGTGACCTCCGAGTCTCAGTAACTGCTTCGTGACTCC 120
QY 121 CATGCTCTTCCACAGCAGACTGAGCCAGTGCCTGCTGAGGAGTTCACCCCTTTGCTACACCTGTC 180
DB 121 CATGCTCTTCCACAGCAGACTGAGCCAGTGCCTGCTGAGGAGTTCACCCCTTTGCTACACCTGTC 180
QY 181 CTGCTGCCTGCTGTGACACTTTAGCTTGGGAGAAATGGAACCCAGATGGAGGACCAAG 240
DB 181 CTGCTGCCTGCTGTGACACTTTAGCTTGGGAGAAATGGAACCCAGATGGAGGACCAAG 240
QY 241 GCACAGGACATTCCTGGAGCAGTGCCTCTCATCCCTCTGCTGGGAGGAGTTCGCTGCTGCTC 300
DB 241 GCACAGGACATTCCTGGAGCAGTGCCTCTCATCCCTCTGCTGGGAGGAGTTCGCTGCTGCTC 300
QY 301 CAAGTGGGACCCACTTGGCTCTCATCCCTCTGCTGGGAGGAGTTCGCTGCTGCTGCTC 360
DB 301 CAAGTGGGACCCACTTGGCTCTCATCCCTCTGCTGGGAGGAGTTCGCTGCTGCTGCTC 360
QY 361 CTCTTGGGGCCCTGCAGAGCCTCTTGGAAACCCAGTTCCTCCACAGGCGAGGACACA 420
DB 361 CTCTTGGGGCCCTGCAGAGCCTCTTGGAAACCCAGTTCCTCCACAGGCGAGGACACA 420
QY 421 GCTCACAAGATCCCAATGCCATCTTCTGAGCTTCCACACCTGCTCCGAGGAAGGTG 480
DB 421 GCTCACAAGATCCCAATGCCATCTTCTGAGCTTCCACACCTGCTCCGAGGAAGGTG 480
QY 481 CGTTTCTGATGCTTGTAGGAGGCTCCACCTCTGCTGAGGCGGGCGCCCAAC 533
DB 481 CGTTTCTGATGCTTGTAGGAGGCTCCACCTCTGCTGAGGCGGGCGCCCAAC 533

RESULT 7
US-08-999-927-1
Sequence 1, Application US/08999927
Patent No. 6099830
GENERAL INFORMATION:
APPLICANT: Kaushansky, Kenneth
TITLE OF INVENTION: Methods of Stimulating Erythropoiesis

TITLE OF INVENTION: Using Hematopoietic Proteins.
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: ZymoGenetics, Inc.
STREET: 1201 Eastlake Avenue East
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98102
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/999,927
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/347,748
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Parker, Gary E
REGISTRATION NUMBER: 31-648
REFERENCE/DOCKET NUMBER: 94-09C2
TELEPHONE: 206-442-6673
TELEFAX: 206-442-6678
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1062 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cdna
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1059
US-08-999-927-1

Query Match 50.3%; Score 526.6; DB 3; Length 1062;
Best Local Similarity 99.2%; Pred. No. 2.1e-119;
Matches 529; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGGAGCTGACTGAATGCTCTCGTGGTCAATGCTTCTCCTAACTCAAGGCTAACGCTG 60
DB 1 ATGGAGCTGACTGAATGCTCTCGTGGTCAATGCTTCTCCTAACTCAAGGCTAACGCTG 60
QY 61 TCCAGCCCGGCTCCTCTGCTTGTGACCTCCGAGTCTCAGTAACTGCTTCGTGACTCC 120
DB 61 TCCAGCCCGGCTCCTCTGCTTGTGACCTCCGAGTCTCAGTAACTGCTTCGTGACTCC 120
QY 121 CATGCTCTTCCACAGCAGACTGAGCCAGTGCCTGCTGAGGAGTTCACCCCTTTGCTACACCTGTC 180
DB 121 CATGCTCTTCCACAGCAGACTGAGCCAGTGCCTGCTGAGGAGTTCACCCCTTTGCTACACCTGTC 180
QY 181 CTGCTGCCTGCTGTGACACTTTAGCTTGGGAGAAATGGAACCCAGATGGAGGACCAAG 240
DB 181 CTGCTGCCTGCTGTGACACTTTAGCTTGGGAGAAATGGAACCCAGATGGAGGACCAAG 240
QY 241 GCACAGGACATTCCTGGAGCAGTGCCTCTCATCCCTCTGCTGGGAGGAGTTCGCTGCTGCTGCTC 300
DB 241 GCACAGGACATTCCTGGAGCAGTGCCTCTCATCCCTCTGCTGGGAGGAGTTCGCTGCTGCTGCTC 300
QY 301 CAAGTGGGACCCACTTGGCTCTCATCCCTCTGCTGGGAGGAGTTCGCTGCTGCTGCTC 360
DB 301 CAAGTGGGACCCACTTGGCTCTCATCCCTCTGCTGGGAGGAGTTCGCTGCTGCTGCTC 360
QY 361 CTCTTGGGGCCCTGCAGAGCCTCTTGGAAACCCAGTTCCTCCACAGGCGAGGACACA 420
DB 361 CTCTTGGGGCCCTGCAGAGCCTCTTGGAAACCCAGTTCCTCCACAGGCGAGGACACA 420


```
; LOCATION: 1..1059
PCT-US95-01829-1

Query Match          50.3%; Score 526.6; DB 5; Length 1062;
Best Local Similarity 99.2%; Pred. No. 2.1e-119;
Matches 529; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGAGCTGACTGAATTCCTCGTGGTCACTGCTTCTCTAACTCAAGGCTAACGCTG 60
Db 1 ATGAGCTGACTGAATTCCTCGTGGTCACTGCTTCTCTAACTCAAGGCTAACGCTG 60
QY 61 TCACAGCCGGCTCCTCTGCTGTGACCTCCGAGTCTCAGTAACTGCTTCGTGACTC 120
Db 61 TCACAGCCGGCTCCTCTGCTGTGACCTCCGAGTCTCAGTAACTGCTTCGTGACTC 120
QY 121 CATGTCTTTCACAGCAGACTGAGCCAGTGCCTCAGAGGTTTCCACCTTTGCCTACACCTGTC 180
Db 121 CATGTCTTTCACAGCAGACTGAGCCAGTGCCTCAGAGGTTTCCACCTTTGCCTACACCTGTC 180
QY 181 CTGCTGCTGCTGCTGGAGCTTTAGCTTGGAGAAATGGAAACCCAGATGGAGGACCAAG 240
Db 181 CTGCTGCTGCTGCTGGAGCTTTAGCTTGGAGAAATGGAAACCCAGATGGAGGACCAAG 240
QY 241 GCACAGGACATTTCTGGGAGCAGTGACCCCTTCTGCTGGAGGAGTGATGGCAGCAGCGGGA 300
Db 241 GCACAGGACATTTCTGGGAGCAGTGACCCCTTCTGCTGGAGGAGTGATGGCAGCAGCGGGA 300
QY 301 CAACTGGGACCATTTGCTGCTCATCCCTCTGCTGGGAGGCTTCTGACAGGTCCTGCTC 360
Db 301 CAACTGGGACCATTTGCTGCTCATCCCTCTGCTGGGAGGCTTCTGACAGGTCCTGCTC 360
QY 361 CTCTTGGGCGCTTGCCTGCTCATCCCTCTGCTGGGAGGCTTCTGACAGGTCCTGCTC 420
Db 361 CTCTTGGGCGCTTGCCTGCTCATCCCTCTGCTGGGAGGCTTCTGACAGGTCCTGCTC 420
QY 421 GCTCACAAGGATCCCAATGCCATCTTCTGAGCTTCCACACCTTCCAGGAGGAAAGGTG 480
Db 421 GCTCACAAGGATCCCAATGCCATCTTCTGAGCTTCCACACCTTCCAGGAGGAAAGGTG 480
QY 481 CGTTTCTGTGCTTTAGAGGCTTCCACCTCTGCTAGCGGGGGGGGGCCCAAC 533
Db 481 CGTTTCTGTGCTTTAGAGGCTTCCACCTCTGCTAGCGGGGGGGGGGGCCCAAC 533

RESULT 10
PCT-US95-14932-3
; Sequence 3, Application PC/TUS9514932
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; APPLICANT: 1201 Eastlake Avenue East
; APPLICANT: Seattle
; APPLICANT: WA
; APPLICANT: USA
; APPLICANT: 98102
; TITLE OF INVENTION: METHOD FOR SECRETING THROMBOPOIETIN
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/14932
; FILING DATE:
```

```
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, Gary E
; REGISTRATION NUMBER: 31-648
; REFERENCE/DOCKET NUMBER: 94-13PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6673
; TELEFAX: 206-442-6678
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1062 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1059
PCT-US95-14932-3

Query Match          50.3%; Score 526.6; DB 5; Length 1062;
Best Local Similarity 99.2%; Pred. No. 2.1e-119;
Matches 529; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGAGCTGACTGAATTCCTCGTGGTCACTGCTTCTCTAACTCAAGGCTAACGCTG 60
Db 1 ATGAGCTGACTGAATTCCTCGTGGTCACTGCTTCTCTAACTCAAGGCTAACGCTG 60
QY 61 TCACAGCCGGCTCCTCTGCTGTGACCTCCGAGTCTCAGTAACTGCTTCGTGACTC 120
Db 61 TCACAGCCGGCTCCTCTGCTGTGACCTCCGAGTCTCAGTAACTGCTTCGTGACTC 120
QY 121 CATGTCTTTCACAGCAGACTGAGCCAGTGCCTCAGAGGTTTCCACCTTTGCCTACACCTGTC 180
Db 121 CATGTCTTTCACAGCAGACTGAGCCAGTGCCTCAGAGGTTTCCACCTTTGCCTACACCTGTC 180
QY 181 CTGCTGCTGCTGCTGGAGCTTTAGCTTGGAGAAATGGAAACCCAGATGGAGGACCAAG 240
Db 181 CTGCTGCTGCTGCTGGAGCTTTAGCTTGGAGAAATGGAAACCCAGATGGAGGACCAAG 240
QY 241 GCACAGGACATTTCTGGGAGCAGTGACCCCTTCTGCTGGAGGAGTGATGGCAGCAGCGGGA 300
Db 241 GCACAGGACATTTCTGGGAGCAGTGACCCCTTCTGCTGGAGGAGTGATGGCAGCAGCGGGA 300
QY 301 CAACTGGGACCATTTGCTGCTCATCCCTCTGCTGGGAGGCTTCTGACAGGTCCTGCTC 360
Db 301 CAACTGGGACCATTTGCTGCTCATCCCTCTGCTGGGAGGCTTCTGACAGGTCCTGCTC 360
QY 361 CTCTTGGGCGCTTGCCTGCTCATCCCTCTGCTGGGAGGCTTCTGACAGGTCCTGCTC 420
Db 361 CTCTTGGGCGCTTGCCTGCTCATCCCTCTGCTGGGAGGCTTCTGACAGGTCCTGCTC 420
QY 421 GCTCACAAGGATCCCAATGCCATCTTCTGAGCTTCCACACCTTCCAGGAGGAAAGGTG 480
Db 421 GCTCACAAGGATCCCAATGCCATCTTCTGAGCTTCCACACCTTCCAGGAGGAAAGGTG 480
QY 481 CGTTTCTGTGCTTTAGAGGCTTCCACCTCTGCTAGCGGGGGGGGGCCCAAC 533
Db 481 CGTTTCTGTGCTTTAGAGGCTTCCACCTCTGCTAGCGGGGGGGGGGGCCCAAC 533

RESULT 11
PCT-US95-16626-3
; Sequence 3, Application PC/TUS9516626
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; APPLICANT: 1201 Eastlake Avenue East
; APPLICANT: Seattle
; APPLICANT: WA
; APPLICANT: USA
; APPLICANT: 98102
; TITLE OF INVENTION: PURIFIED THROMBOPOIETIN AND METHOD OF
```

```

; TITLE OF INVENTION: MAKING IT
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: zymoGenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/16626
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, Gary E
; REGISTRATION NUMBER: 31,648
; REFERENCE/DOCKET NUMBER: 94-11PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6673
; TELEFAX: 206-442-6678
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1062 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1059
; PCT-US95-16626-3

```

Query Match 50.3%; Score 526.6; DB 5; Length 1062;
Best Local Similarity 99.2%; Pred. No. 2.1e-119;
Matches 529; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

1	ATGGAGCTGACTGAATTCCTCTCGTGGTCATGCTTCCTTAACATGCAAGGCTAAACGCTG	60
1	ATGGAGCTGACTGAATTCCTCTCGTGGTCATGCTTCCTTAACATGCAAGGCTAAACGCTG	60
61	TCACGCCGGCTCCTCTGCTTGTGACCTCCGAGTCCCTCAGTAACTGCTTCGTGACTCC	120
61	TCACGCCGGCTCCTCTGCTTGTGACCTCCGAGTCCCTCAGTAACTGCTTCGTGACTCC	120
121	CATGTCTCTTACACAGACTGAGCCAGTGCACAGAGTTACACCTTTTGCTTACACCTGTC	180
121	CATGTCTCTTACACAGACTGAGCCAGTGCACAGAGTTACACCTTTTGCTTACACCTGTC	180
181	CTGTCTGCTCTGTGGACTTACGTGTGGAGAACTGGAACCCAGATGGAGGAGACCAAG	240
181	CTGTCTGCTCTGTGGACTTACGTGTGGAGAACTGGAACCCAGATGGAGGAGACCAAG	240
241	GCACAGGACATTCGGAGCAGTGACCTTCCTGTGGAGGAGTGATGGCAGACACGGGA	300
241	GCACAGGACATTCCTGGAGCAGTGACCTTCCTGTGGAGGAGTGATGGCAGACACGGGA	300
301	CAACTGGGACCCACTGCGCTCTCATCCCTCCTGGGGCAGCTTTCGTGACAGAGTCCGTC	360
301	CAACTGGGACCCACTGCGCTCTCATCCCTCCTGGGGCAGCTTTCGTGACAGAGTCCGTC	360
361	CTCCTTGGGGCCCTGACAGACCTCTCTGGAAACCCAGCTTCCTCCACAGGCGACGACCAC	420
361	CTCCTTGGGGCCCTGCAGAGCCTCCTTGGAAACCCAGCTTCCTCCACAGGCGACGACCAC	420
421	GCTCACAGGATGCCAATGCCATCTTCCTGTAGCTTCCAACTGCTCCGAGGAAAGGTG	480
421	GCTCACAGGATGCCAATGCCATCTTCCTGTAGCTTCCAACTGCTCCGAGGAAAGGTG	480

Qy	481	CGTTTCCTGATGCTTGTTAGGAGGGTCCACCCCTCTCGGTACGGCGGGCGGCCAAC	533
Db	481	CGTTTCCTGATGCTTGTTAGGAGGGTCCACCCCTCTCGGTACGGCGGGCGGCCACC	533
RESULT 12			

```

RESULT 12
US-08-388-779A-1
; Sequence 1, Application US/08388779A
; Patent No. 5596250
; GENERAL INFORMATION:
; APPLICANT: Elliott, Steven G.
; TITLE OF INVENTION: MGDF ANALOGS
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: AMGEN INC.
; STREET: 1840 Dehavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320-1720
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/388,
; FILING DATE: 15-FEB-1995
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Cook Ph.D., Robert R.
; REGISTRATION NUMBER: 31,602
; REFERENCE/DOCKET NUMBER: A-337
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1342 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 36..1094
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 99..1094
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 36..98
US-08-388-779A-1

```

Query Match 50.3%; Score 526.6; DB 1; Length 1342;
Best Local Similarity 99.2%; Pred. NO. 2.2e-119;
Matches 529; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY	1	ATGGAGCTGACTGAATGCTCTCTCGTGGTCATGCTTCCTTAACTCGAAGGCTAACGCTG	60
Db	36	ATGGAGCTGACTGAATGCTCTCTCGTGGTCATGCTTCCTTAACTCGAAGGCTAACGCTG	95
QY	61	TCAGAGCCCGGCTCCCTCTGTTGTGACCTCCAGATCCCTAGTAACTGCTTCGTGACTCC	120
Db	96	TCAGAGCCCGGCTCCCTCTGTTGTGACCTCCAGATCCCTAGTAACTGCTTCGTGACTCC	155
QY	121	CATGTCTCTTCACAGACAGCTGAGCCAGTGCCTCCAGAGGTTACCCCTTGCTTACACTGTC	180
Db	156	CATGTCTCTTCACAGACAGCTGAGCCAGTGCCTCCAGAGGTTACCCCTTGCTTACACTGTC	215
QY	181	CTGCTGCCCTGCTGTGGACTTTAGCTTTGGAGAAATGGAAACCCAGATGGAGAGACCAAG	240
Db	216	CTGCTGCCCTGCTGTGGACTTTAGCTTTGGAGAAATGGAAACCCAGATGGAGAGACCAAG	275

QY 241 GCACAGACATTCCTGGAGCAGTACCCCTTCTGCTGAGGAGTGTGCGCAGCAGCGGGA 300
Db 276 GCACAGACATTCCTGGAGCAGTACCCCTTCTGCTGAGGAGTGTGCGCAGCAGCGGGA 335
QY 301 CAACCTGGAGCCACTTGCCTCTCATCCCTCTGCTGGGGCAGCTTCTGGACAGTCCGCTCTC 360
Db 336 CAACCTGGAGCCACTTGCCTCTCATCCCTCTGCTGGGGCAGCTTCTGGACAGTCCGCTCTC 395
QY 361 CTCCTTGGGGCCCTGCGAGAGCCTCTTGGAAACCCAGCTTCTCCACAGGGCAGGACACCA 420
Db 396 CTCCTTGGGGCCCTGCGAGAGCCTCTTGGAAACCCAGCTTCTCCACAGGGCAGGACACCA 455
QY 421 GCTCACAAAGATCCCAATGCCATCTTCTGAGCTTCCAAACACCTGCTCCGAGGAAAGGTG 480
Db 456 GCTCACAAAGATCCCAATGCCATCTTCTGAGCTTCCAAACACCTGCTCCGAGGAAAGGTG 515
QY 481 CGTTTCTGATGCTTGTAGAGGGTCCACCCCTCTGCTAGCGGGGGGCCCAAC 533
Db 516 CGTTTCTGATGCTTGTAGAGGGTCCACCCCTCTGCTAGCGGGGGGCCCAAC 568

RESULT 13

US-08-591-070A-1

; Sequence 1, Application US/08591070A

; Patent No. 5756083

; GENERAL INFORMATION:

; APPLICANT: Eliott, Steven G.

; TITLE OF INVENTION: MPL LIGAND ANALOGS

; NUMBER OF SEQUENCES: 56

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: AMGEN INC.

; STREET: 1840 Dehavilland Drive

; CITY: Thousand Oaks

; STATE: CA

; COUNTRY: USA

; ZIP: 91320-1789

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/591,070A

; FILING DATE: 09-FEB-1996

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/388,779

; FILING DATE: 15-FEB-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: COOK, ROBERT R

; REGISTRATION NUMBER: 31,602

; REFERENCE/DOCKET NUMBER: A-337B

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1342 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 36..1094

; FEATURE:

; NAME/KEY: mat_peptide

; LOCATION: 99..1094

; FEATURE:

; NAME/KEY: sig_peptide

; LOCATION: 36..98

; US-08-591-070A-1

Query Match

Best Local Similarity 50.3%; Score 526.6; DB 1; Length 1342;

Pred. No. 2.2e-119;

Matches 529; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGAGCTGACTGAATGCTCCTCGTGGTCTGCTTAAGTGAAGCTGAACCTG 60
Db 36 ATGAGCTGACTGAATGCTCCTCGTGGTCTGCTTAAGTGAAGCTGAACCTG 95
QY 61 TCAGACCCGGCTCCTCCTGCTTGTGACCTCGAGTCTCAGTAACTGCTTGTGACTCC 120
Db 96 TCAGACCCGGCTCCTCCTGCTTGTGACCTCGAGTCTCAGTAACTGCTTGTGACTCC 155
QY 121 CATGCTCTTACAGCAGTACGAGCAGTGGCCAGAGGTTTACCCCTTGTGACTGTC 180
Db 156 CATGCTCTTACAGCAGTACGAGCAGTGGCCAGAGGTTTACCCCTTGTGACTGTC 215
QY 181 CTGCTGCTGCTGCTGGAGTCTTAGCTTGGAGAAATGGAACCCAGATGGAGGACCAAG 240
Db 216 CTGCTGCTGCTGCTGGAGTCTTAGCTTGGAGAAATGGAACCCAGATGGAGGACCAAG 275
QY 241 GCACAGACATTCCTGGGAGCAGTACCCCTTCTGCTGGAGGAGTGTGCGCAGCAGCGGGA 300
Db 276 GCACAGACATTCCTGGGAGCAGTACCCCTTCTGCTGGAGGAGTGTGCGCAGCAGCGGGA 335
QY 301 CACTGGGACCCACTTGCCTCTCATCCCTCTGCTGGGAGGAGTGTGCGCAGCAGCGGTC 360
Db 336 CACTGGGACCCACTTGCCTCTCATCCCTCTGCTGGGAGGAGTGTGCGCAGCAGCGGTC 395
QY 361 CTCCTTGGGGCCCTGCGAGAGCCTTGGACCCAGCTTCCACACCTGCTCCACAGGGCAGGACCA 420
Db 396 CTCCTTGGGGCCCTGCGAGAGCCTTGGACCCAGCTTCCACACCTGCTCCACAGGGCAGGACCA 455
QY 421 GCTCACAAAGATCCCAATGCCATCTTCTGAGCTTCCAAACACCTGCTCCGAGGAAAGGTG 480
Db 456 GCTCACAAAGATCCCAATGCCATCTTCTGAGCTTCCAAACACCTGCTCCGAGGAAAGGTG 515
QY 481 CGTTTCTGATGCTTGTAGAGGGTCCACCCCTCTGCTAGCGGGGGGCCCAAC 533
Db 516 CGTTTCTGATGCTTGTAGAGGGTCCACCCCTCTGCTAGCGGGGGGCCCAAC 568

RESULT 14

US-08-413-803-24

; Sequence 24, Application US/08413803

; Patent No. 5766581

; GENERAL INFORMATION:

; APPLICANT: Bartley, Timothy D.

; APPLICANT: Bogenberger, Jakob M.

; APPLICANT: Bosselman, Robert A.

; APPLICANT: Hunt, Pamela

; APPLICANT: Kinstler, Olaf B.

; APPLICANT: Samal, Babru B.

; TITLE OF INVENTION: METHODS FOR TREATING MAMMALS WITH

; TITLE OF INVENTION: MONO-PEGYLATED PROTEINS THAT STIMULATE MEGAKARYOCYTE

; NUMBER OF SEQUENCES: 34

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: AMGEN INC.

; STREET: 1840 Dehavilland Drive

; CITY: Thousand Oaks

; STATE: California

; COUNTRY: US

; ZIP: 91320-1789

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/413,803

; FILING DATE: 30-MAR-1995

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/221,768

; FILING DATE: 31-MAR-1994

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/252,628
; FILING DATE: 31-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE: 12-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/347,780
; FILING DATE: 30-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Cook Ph.D., Robert R.
; REGISTRATION NUMBER: 31,602
; REFERENCE/DOCKET NUMBER: A-290D
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1342 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 36..1097
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 99..1097
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 36..98
; US-08-413-803-24

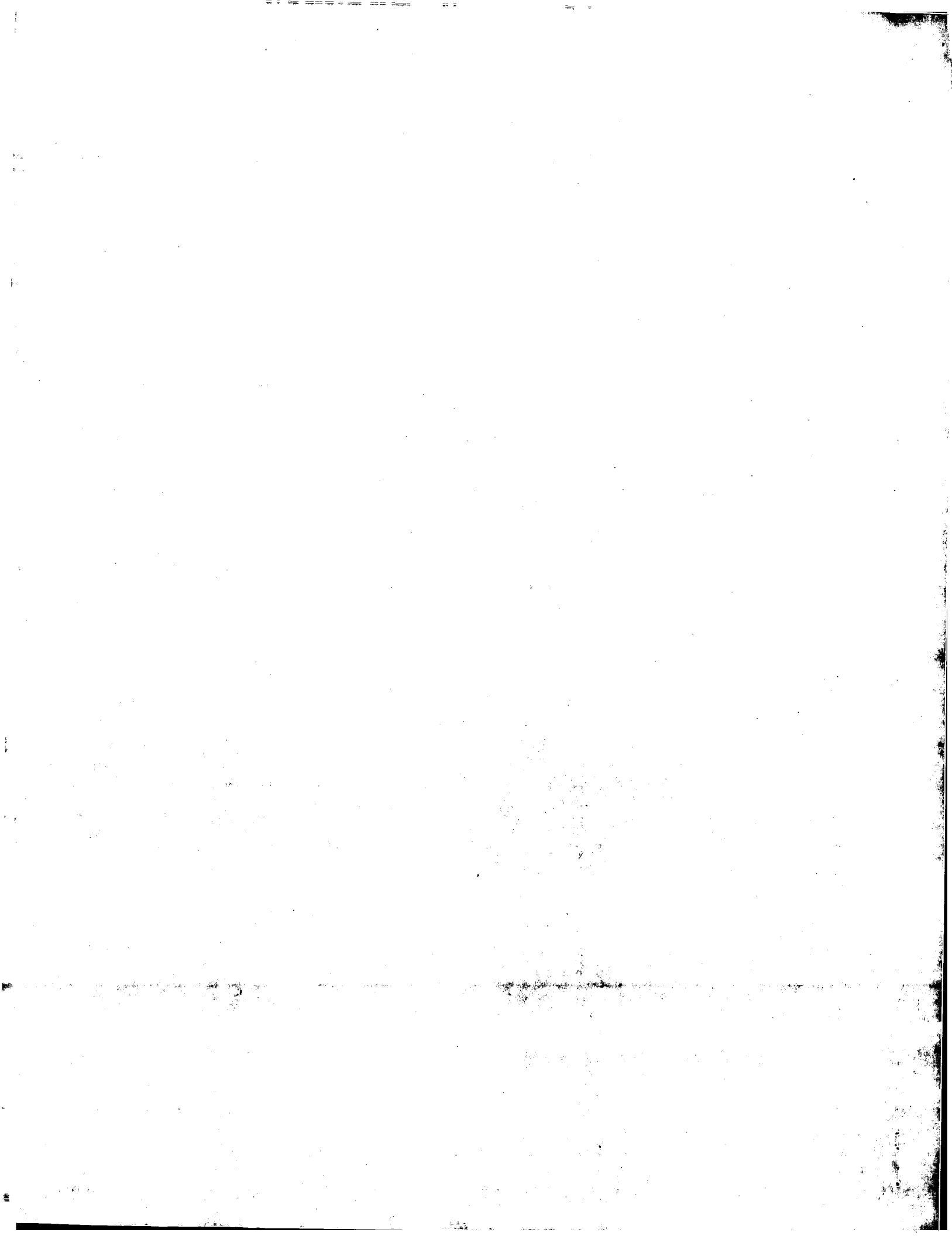
```

Query Match	50.3%	Score 526.6	DB 1	Length 1342
Best Local Similarity	99.2%	Pred. No. 2.2e-119		
Matches 529	Conservative 0	Mismatches 4	Indels 0	Gaps 0
QY 1	ATGGAGCTGACTGAATTGCTCTCTGCTGTCATGCTTCTCTAACTCGAAGGCTAACGCTG	60		
Db				
QY 36	ATGGAGCTGACTGAATTGCTCTCTGCTGTCATGCTTCTCTAACTCGAAGGCTAACGCTG	95		
Db				
QY 61	TCCAGCCCCGGCTCCTCTGCTTTGTGACCTCCGAGTCTCTAGTAAACTGCTTCGTGACTCC	120		
Db				
QY 96	TCCAGCCCCGGCTCCTCTGCTTTGTGACCTCCGAGTCTCTAGTAAACTGCTTCGTGACTCC	155		
Db				
QY 121	CATGTCCTTTCACAGCAGACTGAGCCAGTGCACAGAGTTACCCCTTTGGCTACACCTGTC	180		
Db				
QY 156	CATGTCCTTTCACAGCAGACTGAGCCAGTGCACAGAGTTACCCCTTTGGCTACACCTGTC	215		
Db				
QY 181	CTGCTGCCCTGCTGTGGACTTTAGCTTTGGGAGATGAAACCCAGATGGAGAGACCAAG	240		
Db				
QY 216	CTGCTGCCCTGCTGTGGACTTTAGCTTTGGGAGATGAAACCCAGATGGAGAGACCAAG	275		
Db				
QY 241	GCACAGGACATTCTGGGAGCAGTGACCCCTTCTCTGGAGGGAGTGATGCGACGACGGGA	300		
Db				
QY 276	GCACAGGACATTCTGGGAGCAGTGACCCCTTCTCTGGAGGGAGTGATGCGACGACGGGA	335		
QY 301	CAACTGGGACCCACATGCGCTCTCATCCCTCCTGGGCGACCTTTCTGGACAGTCCGCTCTC	360		
Db				
QY 336	CAACTGGGACCCACATGCGCTCTCATCCCTCCTGGGCGACCTTTCTGGACAGTCCGCTCTC	395		
Db				
QY 361	CTCCTTGGGGCCCTGCAGAGCCTCTTTGGAACCCAGCTTCTCCACAGGGCAGACACACA	420		
Db				
QY 396	CTCCTTGGGGCCCTGCAGAGCCTCTTTGGAACCCAGCTTCTCCACAGGGCAGACACACA	455		
Db				
QY 421	GCTCACAGGATCCCAATGCCATCTTCCTGAGCTTCCAAACCTGCTCCGAGGAAAGGTG	480		
Db				
QY 456	GCTCACAGGATCCCAATGCCATCTTCCTGAGCTTCCAAACCTGCTCCGAGGAAAGGTG	515		
Db				
QY 481	CGTTTCCGTGATGCTGTAGGAGGGTCCACCTCTCGGTACGGGGGCCCAAC	533		
Db				
QY 516	CGTTTCCGTGATGCTGTAGGAGGGTCCACCTCTCGGTACGGGGGCCCAAC	568		
Db				

RESULT 15
 US-08-321-488A-24
 ; Sequence 24, Application US/08321488A
 ; Patent No. 5795569
 ; GENERAL INFORMATION:
 ; APPLICANT: Bartley, Timothy D.
 ; APPLICANT: Bogenberger, Jakob M.
 ; APPLICANT: Bosselman, Robert A.
 ; APPLICANT: Hunt, Pamela
 ; APPLICANT: Kinstler, Olaf B.
 ; APPLICANT: Samal, Babru B.
 ; TITLE OF INVENTION: MONO-PEGYLATED PROTEINS THAT STIMULATE
 ; TITLE OF INVENTION: MEGAKARYOCYTE GROWTH AND DIFFERENTIATION
 ; NUMBER OF SEQUENCES: 27
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: AMGEN INC.
 ; STREET: 1840 DeHavilland Drive
 ; CITY: Thousand Oaks
 ; STATE: California
 ; COUNTRY: US
 ; ZIP: 91320-1789
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; OPERATING SYSTEM: IBM PC compatible
 ; SOFTWARE: Patentin Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/321.488A
 ; FILING DATE: 12-OCT-1994
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/252,628
 ; FILING DATE: 31-MAY-1994
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/221,768
 ; FILING DATE: 31-MAR-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Cook, Robert R.
 ; REGISTRATION NUMBER: 31,602
 ; REFERENCE/DOCKET NUMBER: A-290B
 ; INFORMATION FOR SEQ ID NO: 24:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1342 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: unknown
 ; TOPOLOGY: unknown
 ; MOLECULE TYPE: cDNA
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 36..1097
 ; FEATURE:
 ; NAME/KEY: mat_peptide
 ; LOCATION: 99..1097
 ; FEATURE:
 ; NAME/KEY: sig_peptide
 ; LOCATION: 36..98
 ; US-08-321-488A-24

Query Match	50.3%	Score 526.6	DB 1	Length 1342	
Best Local Similarity	99.2%	pred. No. 2.2e-119			
Matches 529	Conservative	0	Mismatches 4	Indels 0	Gaps 0
Qy 1	ATGGAGTGACTGAATTGCTCTCTCGTGTGCTATGCTTCTCCTAACTGCAAGGCTAACGCTG	60			
Db 36	ATGGAGTGACTGAATTGCTCTCTCGTGTGCTATGCTTCTCCTAACTGCAAGGCTAACGCTG	95			
Qy 61	TCCAGCCCGGCTCCTCTCTGCTGTGTGACCTCCGAGTCCTCAGTAAACTGCTTCGTGACTCC	120			
Db 96	TCCAGCCCGGCTCCTCTCTGCTGTGTGACCTCCGAGTCCTCAGTAAACTGCTTCGTGACTCC	155			
Qy 121	CATGTCTCTTACACGACACTGAGGCAGTGCCCGAGAGTTTACCCCTTTTGCTTACACCTGTC	180			

Search completed: April 5, 2002, 10:34:53
Job time: 4987 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 5, 2002, 09:10:07 ; Search time 2520.15 Seconds
(without alignments)
6853.777 Million cell updates/sec

Title: US-09-680-514-4
Perfect score: 1047
Sequence: 1 ATGAGCTGACTGAATTGCT.....TACGCCACTTGCCAGCCC 1047

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues
Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_om.*
20: em_or.*
21: em_ov.*
22: em_pat.*
23: em_ph.*
24: em_pl.*
25: em_ro.*
26: em_sts.*
27: em_sy.*
28: em_un.*
29: em_vi.*
30: em_htgo_hum.*
31: em_htgo_inv.*
32: em_htgo_rod.*
33: em_htg_hum.*
34: em_htg_inv.*
35: em_htg_rod.*
36: em_htg_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	526.6	50.3	555	22	E11965	E11965 Human cDNA
2	526.6	50.3	1059	6	AR103959	AR103959 Sequence
3	526.6	50.3	1062	6	AR087133	AR087133 Sequence
4	526.6	50.3	1062	6	AR087571	AR087571 Sequence
5	526.6	50.3	1062	6	I49760	I49760 Sequence 3
6	526.6	50.3	1062	9	HUNTHROMB	L36052 Human throm
7	526.6	50.3	1086	6	E16669	E16669 cDNA encodi
8	526.6	50.3	1267	6	E12182	E12182 Human cDNA
9	526.6	50.3	1267	6	E12214	E12214 Human cDNA
10	526.6	50.3	1342	6	AR008878	AR008878 Sequence
11	526.6	50.3	1342	6	AR023468	AR023468 Sequence
12	526.6	50.3	1342	6	AR087573	AR087573 Sequence
13	526.6	50.3	1342	6	I85129	I85129 Sequence 1
14	526.6	50.3	1691	9	D32047	D32047 Homo sapien
15	526.6	50.3	1721	6	E16668	E16668 cDNA encodi
16	526.6	50.3	1721	22	E11961	E11961 Human cDNA
17	526.6	50.3	1795	6	I33525	I33525 Sequence 1
18	526.6	50.3	1795	9	HUMMLCmpl	L33410 Human c-mpl
19	525	50.1	605	6	AR008879	AR008879 Sequence
20	525	50.1	605	6	AR087574	AR087574 Sequence
21	525	50.1	605	6	I85130	I85130 Sequence 3
22	523.4	50.0	1341	9	HSU11025	U11025 Human megak
23	521.8	49.8	1062	9	HSU59493	U59493 Human throm
24	521.8	49.8	1062	9	HSU59494	U59494 Human throm
25	520.4	49.7	525	6	E02574	E02574 DNA encodin
26	514	49.1	525	6	E02575	E02575 DNA encodin
27	512.4	48.9	525	6	E02576	E02576 DNA encodin
28	492.6	47.0	1050	22	E11559	E11559 Human cDNA
29	486.8	46.5	525	6	AR024358	AR024358 Sequence
30	486.8	46.5	525	6	AR091731	AR091731 Sequence
31	486.8	46.5	525	6	E01731	E01731 cDNA encodi
32	486.8	46.5	525	6	E02573	E02573 DNA encodin
33	486.8	46.5	525	6	E07164	E07164 cDNA encodi
34	486.8	46.5	525	6	E05131	E15131 Human mRNA
35	486.8	46.5	525	6	I71150	I71150 Sequence 1
36	486.8	46.5	525	6	I83709	I83709 Sequence 1
37	486.8	46.5	525	22	E09431	E09431 cDNA encodi
38	486.8	46.5	660	9	HUMGCSF	M17706 Human granu
39	486.8	46.5	1498	9	HSGCSFR1	X03655 Human mRNA
40	486.8	46.5	1521	6	E01219	E01219 cDNA encodi
41	486.8	46.5	1525	6	I05333	I05333 Sequence 11
42	486.8	46.5	1525	6	I08669	I08669 Sequence 1
43	486.8	46.5	1525	6	I09205	I09205 Sequence 5
44	485.2	48.3	522	6	E01631	E01631 DNA encodin
45	485.2	46.3	522	6	E08529	E08529 DNA encodin

ALIGNMENTS

RESULT 1
E11965
ID E11965 standard; RNA; HUM; 555 BP.
XX
AC E11965;
XX
SV E11965.1
XX
DT 08-OCT-1997 (Rel. 52, Created)
DT 02-SEP-2000 (Rel. 65, Last updated, Version 2)
XX
DE Human cDNA encoding thrombopoietin.
XX
KW JP 1996228781-A/6.
XX
OS Homo sapiens (human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
XX
RN [1]

BASE COUNT 223 a 375 c 236 g 228 t
ORIGIN

/organism="unknown"

Query Match 50.3%; Score 526.6; DB 6; Length 1062;
Best Local Similarity 99.2%; Pred. No. 1.3e-99;
Matches 529; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGGAGCTGACTGAATGCTCTGCTGTGAGCTCCGAGTCCAGTAACTGCTTGGTACTCC 120
Db 1 ATGGAGCTGACTGAATGCTCTGCTGTGAGCTCCGAGTCCAGTAACTGCTTGGTACTCC 120
QY 61 TCCAGCCCGGCTCTCTGCTGTGAGCTCCGAGTCCAGTAACTGCTTGGTACTCC 180
Db 61 TCCAGCCCGGCTCTCTGCTGTGAGCTCCGAGTCCAGTAACTGCTTGGTACTCC 180
QY 121 CATGCTCTTACAGCAGACTGAGCAGTCCGAGTCCAGTAACTGCTTGGTACTCC 120
Db 121 CATGCTCTTACAGCAGACTGAGCAGTCCGAGTCCAGTAACTGCTTGGTACTCC 120
QY 181 CTGCTGCTCTGCTGCTCTGCTGTGAGCTCCGAGTCCAGTAACTGCTTGGTACTCC 180
Db 181 CTGCTGCTCTGCTGCTCTGCTGTGAGCTCCGAGTCCAGTAACTGCTTGGTACTCC 180
QY 241 GCACAGGACATTTCTGGGAGCAGTCCGAGTCCAGTAACTGCTTGGTACTCC 240
Db 241 GCACAGGACATTTCTGGGAGCAGTCCGAGTCCAGTAACTGCTTGGTACTCC 240
QY 241 GCACAGGACATTTCTGGGAGCAGTCCGAGTCCAGTAACTGCTTGGTACTCC 240
Db 241 GCACAGGACATTTCTGGGAGCAGTCCGAGTCCAGTAACTGCTTGGTACTCC 240
QY 301 CACTGGGACCACTTCTGCTCTGCTGTGAGCTCCGAGTCCAGTAACTGCTTGGTACTCC 300
Db 301 CACTGGGACCACTTCTGCTCTGCTGTGAGCTCCGAGTCCAGTAACTGCTTGGTACTCC 300
QY 361 CTCTTGGGCGCTCGACAGCTCTCTGCTGTGAGCTCCGAGTCCAGTAACTGCTTGGTACTCC 360
Db 361 CTCTTGGGCGCTCGACAGCTCTCTGCTGTGAGCTCCGAGTCCAGTAACTGCTTGGTACTCC 360
QY 421 GCTCAAGGATCCCAATGCTCTGCTGTGAGCTCCGAGTCCAGTAACTGCTTGGTACTCC 420
Db 421 GCTCAAGGATCCCAATGCTCTGCTGTGAGCTCCGAGTCCAGTAACTGCTTGGTACTCC 420
QY 481 CGTTTCTGATGCTGTAGGAGGTCCACCTCTGCTGTGAGCTCCGAGTCCAGTAACTGCTTGGTACTCC 480
Db 481 CGTTTCTGATGCTGTAGGAGGTCCACCTCTGCTGTGAGCTCCGAGTCCAGTAACTGCTTGGTACTCC 480

RESULT 6
LOCUS HUNTHROMB 1062 bp mRNA PRI 22-DEC-1994
DEFINITION Human thrombopoietin mRNA, complete cds.
ACCESSION L36052
VERSION L36052.1 GI:533216
KEYWORDS Homo sapiens cDNA to mRNA.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1062)
AUTHORS Foster,D.C., Sprecher,C.A., Grant,F.J., Kramer,J.M., Kuijper,J.L.,
Holly,R.D., Whitmore,T.E., Heipel,M.D., Bell,L.A.N., Ching,A.F.,
McGrane,V., Hart,C., O'Hara,P.J., and Lok,S.
TITLE Human thrombopoietin: gene structure, cDNA sequence, expression,
and chromosomal localization
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 91 (26), 13023-13027 (1994)
MEDLINE 95108091
FEATURES
Source Location/Qualifiers
1..1062
/organism="Homo sapiens"
/db_xref="taxon:9606"
CDS 1..1062
/note="putative"
/codon_start=1
/product="thrombopoietin"

/protein_id="AAC37566.1"
/db_xref="GI:533217"

/translation="MELTELLLVMLLTARLTLSSPAPACDLRVLSKLLRDSHVLH
SRLSQCEVHPLETPVLLPAVDFSLGEMKTMETKAQDILGAVTLILGCVMAARGOL
GPYCSLLSGLSGVRLGALSLGLTQLPQGRTHAKDPNATLFLFOHLLRQKV
RFLMLVGGSTLCVRRAPPTAVPSRTSLVLTNLEPNRTSGLETTASRTTGSGL
LKWOQGFRAKIPGLNLSLDOIPCYLNRHELLNGRLGFPSPRTTLGAPDISS
GTSDDTSLPNLPQGYSPSPTHPTQYTLFLPLPTLPVQLHPLDPSAPTPTP
TSPLLNTSVTHSNLSOEG"

BASE COUNT 223 a 375 c 236 g 228 t
ORIGIN

Query Match 50.3%; Score 526.6; DB 9; Length 1062;
Best Local Similarity 99.2%; Pred. No. 1.3e-99;
Matches 529; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGGAGCTGACTGAATGCTCTGCTGTGAGCTCCGAGTCCAGTAACTGCTTGGTACTCC 60
Db 1 ATGGAGCTGACTGAATGCTCTGCTGTGAGCTCCGAGTCCAGTAACTGCTTGGTACTCC 60
QY 61 TCCAGCCCGGCTCTCTGCTGTGAGCTCCGAGTCCAGTAACTGCTTGGTACTCC 60
Db 61 TCCAGCCCGGCTCTCTGCTGTGAGCTCCGAGTCCAGTAACTGCTTGGTACTCC 60
QY 121 CATGCTCTTACAGCAGACTGAGCAGTCCGAGTCCAGTAACTGCTTGGTACTCC 120
Db 121 CATGCTCTTACAGCAGACTGAGCAGTCCGAGTCCAGTAACTGCTTGGTACTCC 120
QY 181 CTGCTGCTCTGCTGCTCTGCTGTGAGCTCCGAGTCCAGTAACTGCTTGGTACTCC 180
Db 181 CTGCTGCTCTGCTGCTCTGCTGTGAGCTCCGAGTCCAGTAACTGCTTGGTACTCC 180
QY 241 GCACAGGACATTTCTGGGAGCAGTCCGAGTCCAGTAACTGCTTGGTACTCC 240
Db 241 GCACAGGACATTTCTGGGAGCAGTCCGAGTCCAGTAACTGCTTGGTACTCC 240
QY 301 CACTGGGACCACTTCTGCTCTGCTGTGAGCTCCGAGTCCAGTAACTGCTTGGTACTCC 300
Db 301 CACTGGGACCACTTCTGCTCTGCTGTGAGCTCCGAGTCCAGTAACTGCTTGGTACTCC 300
QY 361 CTCTTGGGCGCTCGACAGCTCTCTGCTGTGAGCTCCGAGTCCAGTAACTGCTTGGTACTCC 360
Db 361 CTCTTGGGCGCTCGACAGCTCTCTGCTGTGAGCTCCGAGTCCAGTAACTGCTTGGTACTCC 360
QY 421 GCTCAAGGATCCCAATGCTCTGCTGTGAGCTCCGAGTCCAGTAACTGCTTGGTACTCC 420
Db 421 GCTCAAGGATCCCAATGCTCTGCTGTGAGCTCCGAGTCCAGTAACTGCTTGGTACTCC 420
QY 481 CGTTTCTGATGCTGTAGGAGGTCCACCTCTGCTGTGAGCTCCGAGTCCAGTAACTGCTTGGTACTCC 480
Db 481 CGTTTCTGATGCTGTAGGAGGTCCACCTCTGCTGTGAGCTCCGAGTCCAGTAACTGCTTGGTACTCC 480

RESULT 7
LOCUS E16669 1086 bp DNA PAT 28-JUL-1999
DEFINITION cDNA encoding thrombopoietin.
ACCESSION E16669
VERSION E16669.1 GI:5711352
KEYWORDS JP 1998212243-A/3.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1086)
AUTHORS Kuroda,K.
TITLE COMPOSITION CONTAINING THROMBOPOIETIN PROTEIN AND USED FOR
STIMULATING RECOVERY OF HEMOSTATIC THROMBUS-FORMING ABILITY OF
PATENT: JP 1998212243-A 3 11-AUG-1998;
JOURNAL KIRIN BREWERY CO LTD
COMMENT OS Homo sapiens (human)
PN JP 1998212243-A/3

DEFINITION	Human cDNA encoding thrombopoietin.
ACCESSION	E12182
VERSION	E12182.1 GI:3251016
KEYWORDS	JP 1996277296-A/2.
SOURCE	Homo sapiens.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 (bases 1 to 1267) Miyazaki, H., Kato, H., Ogami, K., Iwamatsu, A., Akahori, H., Kuroki, R., Shimizu, T. and Muto, T.
TITLE	PROTEIN HAVING TPO ACTIVITY
JOURNAL	Patent: JP 1996277296-A 2 22-OCT-1996;
COMMENT	KIRIN BREWERY CO LTD OS Homo sapiens (human) PN JP 1996277296-A/2 PD 22-OCT-1996 PF 14-FEB-1995 JP 1995161363 PR 14-FEB-1994 JP 94P 39090, 25-MAR-1994 JP 94P 79842, PR 01-JUN-1994 JP 94P 155126, 15-JUN-1994 JP 94P 167328, PR 17-AUG-1994 JP 94P 227159, 17-AUG-1994 JP 94P 193169, PR 01-NOV-1994 JP 94P 304167, 01-DEC-1994 JP 94P 298669, PR 28-DEC-1994 JP 94P 341200 PI MIYAZAKI HIROSHI, KATO HISASHI, OGAMI KINYA, IWAMATSU AKIHIKO, PI AKAHORI HIRONORI, KUROKI RYOTA, SHIMIZU TOSHIYUKI, PI MUTO TAKANORI PC C07K14/52, C07H21/04, C12N15/09, C12P21/02, //A61K38/00, (C12P21/02,

[illegible]

RESULT
E12182

```

Db 325 CAACGGAGCCCACTTGCCTCTCATCCCTCCCTGGGGAGCTTCTGGACAGGTCCTGTC 384
Qy 361 CTCCTGGGGCCCTGCAGAGCTCTCTTGAACCCAGCTTCTCCACAGGCGAGCACACA 420
Db 385 CTCCTGGGGCCCTGCAGAGCTCTTGAACCCAGCTTCTCCACAGGCGAGCACACA 444
Qy 421 GCTCACAAGGATCCCAATGCCATCTCTGAGCTTCCCAACACCTGCTCCGAGGAAGGTG 480
Db 445 GCTCACAAGGATCCCAATGCCATCTCTGAGCTTCCCAACACCTGCTCCGAGGAAGGTG 504
Qy 481 CGTTTCCTGATGCTTGTAGGAGGTCCACCTCTGGTACGGGGCGGCCCAAC 533
Db 505 CGTTTCCTGATGCTTGTAGGAGGTCCACCTCTGGTACGGGGCGGCCCAAC 557

RESULT 9
E12214
LOCUS Human cDNA encoding thrombopoietin. PAT 24-JUN-1998
DEFINITION E12214
ACCESSION E12214
VERSION E12214.1 GI:3251048
KEYWORDS JP 1996291196-A/1.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1267)
AUTHORS Miyazaki,H., Kato,H., Ogami,K., Iwamatsu,A., Akahori,H., Kuroki,R.,
Shimizu,T. and Muto,T.
TITLE PROTEIN HAVING TPO ACTIVITY
JOURNAL Patent: JP 1996291196-A 1 05-NOV-1996;
KIRIN BREWERY CO LTD
COMMENT OS Homo sapiens (human)
PN JP 1996291196-A/1
PD 05-NOV-1996
PF 14-FEB-1995 JP 1995355052
PR 14-FEB-1994 JP 94P 39090, 25-MAR-1994 JP 94P 79842, PR
01-JUN-1994 JP 94P 155126, 15-JUN-1994 JP 94P 167328, PR
17-AUG-1994 JP 94P 227159, 17-AUG-1994 JP 94P 193169, PR
01-NOV-1994 JP 94P 304167, 01-DEC-1994 JP 94P 298669, PR
28-DEC-1994 JP 94P 341200
PI MIYAZAKI HIROSHI, KATO HISASHI, OGAMI KINYA, IWAMATSU AKIHIKO,
PI AKAHORI HIRONORI, KUROKI RYOTA, SHIMIZU TOSHIYUKI, PI MUTO
TAKANORI
PC C07K14/53,C07K1/22,C12N15/09//A61K38/00,C12P21/08; CC
Strandedness: Double;
CC topology: Linear;
FH Key Location/Qualifiers
FT source 1..1267
FT CDS /organism='Homo sapiens'
FT /tissue_type='liver'
FT 25..1086
FT sig_peptide 25..87
FT mat_peptide 88..1083
FT /product='thrombopoietin'
FT /product='thrombopoietin'.
FEATURES source
1..1267
Location/Qualifiers
/organism='Homo sapiens'
/db_xref='taxon:9606'
BASE COUNT 281 a 436 c 279 g 271 t
ORIGIN

Query Match 50.3%; Score 526.6; DB 6; Length 1267;
Best Local Similarity 99.2%; Pred. No. 1.3e-99;
Matches 529; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ATGGAGCTGACTGAATTCCTCGTGGTCAATCTCTCTCACTGCAAGGCTAACGCTG 60
Db 25 ATGGAGCTGACTGAATTCCTCGTGGTCAATCTCTCTCACTGCAAGGCTAACGCTG 84

```

```

Qy 61 TCCAGCCCGGCTCCTCTGCTGTGACCTCCGAGTCCCTCAGTAAACTGCTTGGTACTCC 120
Db 85 TCCAGCCCGGCTCCTCTGCTGTGACCTCCGAGTCCCTCAGTAAACTGCTTGGTACTCC 144
Qy 121 CATGCTCTTACAGCAGACTGAGCCAGTCCGAGAGGTTACCCCTTGCCTACACCTGTC 180
Db 145 CATGCTCTTACAGCAGACTGAGCCAGTCCGAGAGGTTACCCCTTGCCTACACCTGTC 204
Qy 181 CTGCTGCCCTGCTGGACTTTAGCTTTGGGAGAAATGGAACCCAGATGGAGGAGCAAG 240
Db 205 CTGCTGCCCTGCTGGACTTTAGCTTTGGGAGAAATGGAACCCAGATGGAGGAGCAAG 264
Qy 241 GCACAGGACATTTCTGGGAGCAGTGCACCTTCTGCTGGAGGAGTGATGGCAGCACGGGA 300
Db 265 GCACAGGACATTTCTGGGAGCAGTGCACCTTCTGCTGGAGGAGTGATGGCAGCACGGGA 324
Qy 301 CAACTGGGACCCACTTGCCTCTCATCCTCTCTGGGAGAGCTTTCTGACAGGTCCTGTC 360
Db 325 CAACTGGGACCCACTTGCCTCTCATCCTCTCTGGGAGAGCTTTCTGACAGGTCCTGTC 384
Qy 361 CTCCTTGGGGCCCTGACAGCCCTCTTGAACCCAGCTTCTCCACAGGCGAGGACACA 420
Db 385 CTCCTTGGGGCCCTGACAGCCCTCTTGAACCCAGCTTCTCCACAGGCGAGGACACA 444
Qy 421 GCTCACAAGGATCCCAATGCCATCTTCTGAGCTTCCCAACACCTGCTCCGAGGAAAGTG 480
Db 445 GCTCACAAGGATCCCAATGCCATCTTCTGAGCTTCCCAACACCTGCTCCGAGGAAAGTG 504
Qy 481 CGTTTCCTGATGCTTGTAGGAGGTCCACCTCTGCTGACGGCGGCCCAAC 533
Db 505 CGTTTCCTGATGCTTGTAGGAGGTCCACCTCTGCTGACGGCGGCCCAAC 557

RESULT 10
AR008878
LOCUS AR008878 1342 bp DNA
DEFINITION Sequence 1 from patent US 5756083.
ACCESSION AR008878
VERSION AR008878.1 GI:3967683
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1342)
AUTHORS Elliott,S.G.
TITLE Mpl ligand analogs
JOURNAL Patent: US 5756083-A 1 26-MAY-1998;
FEATURES Location/Qualifiers
1..1342
Source
BASE COUNT 299 a 454 c 293 g 296 t
ORIGIN

Query Match 50.3%; Score 526.6; DB 6; Length 1342;
Best Local Similarity 99.2%; Pred. No. 1.3e-99;
Matches 529; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ATGGAGCTGACTGAATTCCTCGTGGTCAATCTCTCTCACTGCAAGGCTAACGCTG 60
Db 36 ATGGAGCTGACTGAATTCCTCGTGGTCAATCTCTCTCACTGCAAGGCTAACGCTG 95
Qy 61 TCCAGCCCGGCTCCTCTGCTGTGACCTCCGAGTCCCTCAGTAAACTGCTTGGTACTCC 120
Db 96 TCCAGCCCGGCTCCTCTGCTGTGACCTCCGAGTCCCTCAGTAAACTGCTTGGTACTCC 155
Qy 121 CATGCTCTTACAGCAGACTGAGCCAGTGCACAGAGGTTACCCCTTGCCTACACCTGTC 180
Db 156 CATGCTCTTACAGCAGACTGAGCCAGTGCACAGAGGTTACCCCTTGCCTACACCTGTC 215
Qy 181 CTGCTGCCCTGCTGGACTTTAGCTTTGGGAGAAATGGAACCCAGATGGAGGAGCAAG 240
Db 216 CTGCTGCCCTGCTGGACTTTAGCTTTGGGAGAAATGGAACCCAGATGGAGGAGCAAG 275

```



```
|||||
Db 522 GCTCACAGGATCCCAATGCCATCTTCTGAGCTTCCAACACCTGCTCCGAGGAAAGGTG 581
Qy 481 CGTTTCTGTGCTGTAGGAGGTCCACCTCTGCGGTACGGCGGGCCCAAC 533
Db 582 CGTTTCTGTGCTGTAGGAGGTCCACCTCTGCGGTACGGCGGGCCCAAC 634

RESULT 15
E16668
LOCUS E16668 1721 bp DNA PAT 28-JUL-1999
DEFINITION cDNA encoding thrombopoietin.
ACCESSION E16668
VERSION E16668.1 GI:57111351
KEYWORDS JP 1998212243-A/2.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Kuroda,K
1 (bases 1 to 1721)
COMPOSITION CONTAINING THROMBOPOIETIN PROTEIN AND USED FOR
STIMULATING RECOVERY OF HEMOSTATIC THROMBUS-FORMING ABILITY OF
PATENT: JP 1998212243-A 2 11-AUG-1998;
KIRIN BREWERY CO LTD
OS Homo sapiens (human)
PN JP 1998212243-A/2
PD 11-AUG-1998
PF 31-JAN-1997 JP 1997019549
PI KURODA KENJI
PC A61K38/00,A01N1/02,C07K14/52//C12N15/09,C12P21/02,(C12P21/02,
C12R1:19);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
FH key Location/Qualifiers
FH 1..1721
FT source /organism='Homo sapiens'
FT /tissue_type='Liver'
FT 1..101
FT 5'UTR
FT sig_peptide 102..164
FT /product='signal peptide of thrombopoietin' FT
mat_peptide 165..1160
FT /product='thrombopoietin'
FT CDS 102..1163
FT /product='thrombopoietin'
FT 1164..1721.
FT 3'UTR Location/Qualifiers
FT 1..1721
FT /organism='Homo sapiens'
FT /db_xref='taxon:9606'
BASE COUNT 413 a 542 c 371 g 395 t
ORIGIN
```

```
Query Match 50.3%; Score 526.6; DB 6; Length 1721;
Best Local Similarity 99.2%; Pred. No. 1.2e-99;
Matches 529; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ATGGAGCTGACTGAATGCTCCGTGTCATGCTTCTCCTTAAGTGAAGGCTAACGCTG 60
Db 102 ATGGAGCTGACTGAATGCTCCGTGTCATGCTTCTCCTTAAGTGAAGGCTAACGCTG 161
Qy 61 TCCAGCCCGGCTCCTCTGCTGTGACCTCCGAGTCTCAGTAACTGCTTGGTGACTCC 120
Db 162 TCCAGCCCGGCTCCTCTGCTGTGACCTCCGAGTCTCAGTAACTGCTTGGTGACTCC 221
Qy 121 CATGTCCTTCACAGCAGCTGAGCCAGTGCACAGGTTTCAACCTTTGCGTTACACCTGTC 180
Db 222 CATGTCCTTCACAGCAGCTGAGCCAGTGCACAGGTTTCAACCTTTGCGTTACACCTGTC 281
```

```
Qy 181 CTGCTGCTGCTGTGTGACTTTAGCTTGGGAGATGGAACCCAGATGGAGGAGACCAAG 240
Db 282 CTGCTGCTGCTGTGTGACTTTAGCTTGGGAGATGGAACCCAGATGGAGGAGACCAAG 341
Qy 241 GCACAGGACATTCCTGGGAGCAGTACCCCTTCTGCTGGAGGAGTGTGGCAGCAGGGGA 300
Db 342 GCACAGGACATTCCTGGGAGCAGTACCCCTTCTGCTGGAGGAGTGTGGCAGCAGGGGA 401
Qy 301 CAACTGGGACCCACTTGCCTCTCATCCCTCTCTGGGCGAGCTTTCTTGACAGGTCCGTCTC 360
Db 402 CAACTGGGACCCACTTGCCTCTCATCCCTCTCTGGGCGAGCTTTCTTGACAGGTCCGTCTC 461
Qy 361 CTCCTTGGGGCCCTGCAGAGCCTCCTTTGGAACCCAGCTTCTCCACAGGGCAGGACACA 420
Db 462 CTCCTTGGGGCCCTGCAGAGCCTCCTTTGGAACCCAGCTTCTCCACAGGGCAGGACACA 521
Qy 421 GCTCACAAGATCCCAATGCCATCTTCTGAGCTTCCAACACCTGCTCCGAGGAAAGGTG 480
Db 522 GCTCACAAGATCCCAATGCCATCTTCTGAGCTTCCAACACCTGCTCCGAGGAAAGGTG 581
Qy 481 CGTTTCTGTGCTGTAGGAGGTCCACCTCTGCGTACGGCGGGCGCCCAAC 533
Db 582 CGTTTCTGTGCTGTAGGAGGTCCACCTCTGCGTACGGCGGGCGCCCAAC 634
```

Search completed: April 5, 2002, 10:29:29
Job time: 4762 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 5, 2002, 09:10:41 ; Search time 188.75 Seconds
(without alignments)
4755.600 Million cell updates/sec

Title: US-09-680-514-4
Perfect score: 1047
Sequence: 1 ATGGAGCTGACTGAATTGCT.....TACGCCACCTTGCCAGGCC 1047

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 930621 seqs, 428662619 residues
Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_1101.*
1: /SIDS2/gcgdata/geneseq/geneseq/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1047	100.0	1047	17	AA141786
2	993	94.8	1083	17	AA141787
3	987.4	94.3	1095	17	AA141788
4	526.6	50.3	861	16	AAQ99552
5	526.6	50.3	1062	16	AA103941
6	526.6	50.3	1062	16	AA104050
7	526.6	50.3	1062	17	AA134852
8	526.6	50.3	1062	17	AA137383
9	526.6	50.3	1062	17	AA132591
10	526.6	50.3	1062	18	AA185555
11	526.6	50.3	1062	19	AA121696

12	526.6	50.3	1062	21	AAA51991	Human thrombopoietin
13	526.6	50.3	1086	17	AA147958	Human thrombopoietin
14	526.6	50.3	1342	16	AAQ99704	Human MGDF-1 and M
15	526.6	50.3	1342	17	AA136657	Native human mpl 1
16	526.6	50.3	1342	17	AA110025	Human MGDF-1/2 CDN
17	526.6	50.3	1721	16	AAQ99554	Thrombopoietin cod
18	526.6	50.3	1721	21	AA240191	Human wild type th
19	526.6	50.3	1774	21	AA147793	Human thrombopoiet
20	526.6	50.3	1795	16	AAQ94107	HML cDNA. Homo sa
21	526.6	50.3	1795	18	AA164318	Human thrombopoiet
22	525	50.1	600	17	AA136658	Truncated human mp
23	525	50.1	605	20	AA132813	Human truncated mp
24	525	50.1	1267	16	AAQ99553	Thrombopoietin cod
25	514	49.1	525	11	AAQ04482	Plasmid PASN6 enco
26	510.8	48.8	525	11	AAQ04484	Plasmid PASN145 en
27	506	48.3	525	11	AAQ04481	Plasmid PAS28 enco
28	503.6	48.1	1342	17	AA133933	Human megakaryocyt
29	503.6	48.1	1342	19	AA129068	Nucleotide sequenc
30	503.6	48.1	1342	20	AA132812	Human mpl ligand e
31	492.6	47.0	1050	17	AA141901	Human thrombopoiet
32	486.8	46.5	525	9	AA180947	G-CSF gene isolate
33	486.8	46.5	1520	8	AA170223	Plasmid pBRV2 inse
34	486.8	46.5	1520	8	AA171320	Sequence encoding
35	486.8	46.5	1521	7	AA160937	Plasmid pBRV2 inse
36	486.8	46.5	1525	9	AA181478	Sequence encoding
37	486.8	46.5	1525	10	AA191086	Plasmid pP12 contg
38	485.2	46.3	1415	8	AA171089	Sequence of human
39	479.8	45.8	741	17	AA141907	Human thrombopoiet
40	478	45.7	861	17	AA141902	Human thrombopoiet
41	478	45.7	861	17	AA15486	Megakaryocyte diff
42	478	45.7	946	17	AA129735	Human platelet pro
43	478	45.7	946	21	AA244174	Human platelet-pro
44	478	45.7	946	21	AA244179	Human platelet-pro
45	478	45.7	1164	16	AAQ99705	Human MGDF-3. Hom

ALIGNMENTS

RESULT 1
AA141786
ID AAT41786 standard; DNA; 1047 BP.
XX
AC AAT41786;
XX
DT 01-JUL-1997 (first entry)
XX
DE Fusion peptide #1 having G-CSF and TPO activity.
XX
Fusion protein; human granulocyte colony stimulating factor; hg-CSF;
thrombopoietin; TPO; spacer peptide; blood platelet production;
leukocyte production; anaemia; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT sig_peptide 1..63
FT mat_peptide 64..1047
FT /*tag= a
FT /*tag= b
XX
XX WO9634016-A1.
XX
XX 31-OCT-1996.
XX
XX 26-APR-1996; 96WO-JP01157.
XX
XX 26-APR-1995; 95JP-0102625.
XX
XX (KYOW) KYOWA HAKKO KOGYO KK.
XX
XX Anazawa H, Konishi N, Shiotsu Y, Tamaoki T, Terasaki Y;
PI Uchida K, Yamasaki M, Yamashita K, Yokoi H;

(TPO) activity, opt. via a spacer peptide. Peptides derived from these by deletion, insertion or substitution of one or more amino acid residues are included within the scope of the invention. The fusion peptides stimulate blood platelet and leukocyte production and are useful in the treatment of anaemia.

SQ Sequence 1083 BP; 197 A; 355 C; 305 G; 226 T; 0 other;

Query Match	94.8%	Score 993;	DB 17;	Length 1083;
Best Local Similarity	96.2%	Pred. NO. 1.9e-228;		
Matches 1042;	Conservative	0;	Mismatches 5;	Indels 36;
				Gaps 1;

Qy	1	ATGAGCTGACTGAATTCCTCTCTCGTGGTTCATGCTTCTCTTAAGCTTAAAGCTGCTG	60
Db	1	atgagctgactgaatfgctctctctggtcatgcttctcttaactcaaggtcaagctaaagctg	60
Qy	61	TCCAGCCGGCTCCTCTGTTGTGACCTCCGAGTCTCTCAGTAACTGCTTCGTGACTCC	120
Db	61	tccagccggctcctctgcttggtagctccgagtcctcaagtaaaactgcttcgtgactcc	120
Qy	121	CATGTCCTTTCACAGCAGCTGAGCCAGTGCCTCAGAGAGTTTACCCCTTTGCCCTACACCTGTC	180
Db	121	catgtcttcacagcagactgagccagtgccagaggttcaccccttgcctacacctgtc	180
Qy	181	CTGCTGCTGCTGTGGACTTTAGCTTCGGAGATGGAACACCCAGATGGAGGAGACCAAG	240
Db	181	ctgcctgcctgctgk9gactttgacttcgagtcctcaagtaaaactgcttcgtgactcc	240
Qy	241	GCACAGGACATTCCTGGGAGCAGTGAACCTTCTGCTGGAGGAGTGTATGCCAGCAGCGGGA	300
Db	241	gcacaggacattctg9gagcagtgaacctctctg9gagggagtgatggcagcac9ggga	300
Qy	301	CAACTGGACCCACTTGCCTCTCATTCCTCCTGGGCGAGCTTTCGTGGACAGGTCCTGTC	360
Db	301	caactgg9aaccaactgctctctcatccctctgg9cagcttctg9aacggtctcgtctc	360
Qy	361	CTCTTTGGGGCCCTGTCAGAGCCTCTTGTGAACCCAGCTTCTCCACAGGCGAGGACCA	420
Db	361	ctctctggggccctg9cagcctctctg9aacccagcttctccacagggcaggaccaca	420
Qy	421	GCTCACAAAGNATCCCATGCAATCTTCTGAGCTTCCACACACTGCTCCGAGGAAAGGTG	480
Db	421	gctcacaaagatcccaatgcatctctctgagcttccaacacctgctccgaggaagaggtg	480
Qy	481	CGTTTCTGTGCTGTGTAGGAGGTCACACCTCTGCTGATCGGCGG-----525	525
Db	481	cgtttctgtatgctgttaggaggggtccacccctctgctcagggggttcctggaggt	540
Qy	526	-----GGCCCAACATATCGCGCCTCGAGTCTACACAGAGCTTC	564
Db	541	ggttcc9gagggggtctagagcaccaacatafcg9cctc9agctctcacacagagcttc	600
Qy	565	CTTTTAAAAAGCTTAGAGCAAGTACGGAAGATCCAGGGCGATCGCGAGCTTCAGAGGAG	624
Db	601	cttttaa9agcttag9agcaagtg9aggaatccag99gagtg9gcagc9ctccc99ag	660
Qy	625	AAGCTGTGTCACCTACAAAGCTGTGCCACCCCGAGGAGTGTGTGCTCGACACTCT	684
Db	661	aagctgtgtgccacctacaagctgtgccaccc9ag9agctggtgtgctc9gacactct	720
Qy	685	CTGGGCATCCCTTGGGCTCCCTGAGCAGAGTGTGCCAGCCAGCCCTTGAGCTGGCAGGC	744
Db	721	ctgg9catccctctgg9ctccctctg9agctgtcccacgcag9ccctgagcttggcag9c	780
Qy	745	TGCTGTAGCCCAACTTCCATAGCGGCTTTTCTCTTACCCAGGGGCTCTCTGACGGCCCTGGAA	804
Db	781	tgctttagccaaactccatag9gccttttctctacac9gg9ctctg9ag9ccttg9aa	840
Qy	805	GGGATCTCCCGAGTTGGGTGCCACTTGGACACACTGCAGCTGACGTGCGCGACTTT	864
Db	841	gg9atctccccaggttgggtcccaacttg9acacactg9acactg9actg9ac9tgc9acttt	900

Qy	865	GCCACCACCATCTGTGGCAGCAGATGGGAAGAACTGGGAATGGCCCTTGCCCTGCAGCCCAACC	924
Db	901	gccaccaccatctggcagcagatgggaagaactgggaatggccctgcccctgcagcccaacc	960
Qy	925	CAGGGTGCCATGCGGGGCTTCGGCTCTGCTTTCCAGCGCCGGGAGAGAGGGGTCTTAGTT	984
Db	961	cagggtgcattgcgcggccttcgcctctgtttccacgcgcgggcagggggtcctagt	1020
Qy	985	GCCTTCCCATCTGCAGAGACTTCCTGGAGGTGTCGTACCGGGTTCACGCCACCTTGCCCGAG	1044
Db	1021	gcctcccatctgcagagcttcctcctggagggtcgtaccgcgcttctacgcacacttgccacg	1080
Qy	1045	CCC 1047	
Db	1081	ccc 1083	
RESULT 3			
AAT41788			
ID	AAT41788 standard; DNA; 1095 BP.		
XX	XX		
AC	AAT41788;		
XX	XX		
DT	01-JUL-1997 (first entry)		
XX	XX		
DE	Fusion peptide #3 having G-CSF and TPO activity.		

DE Fusion peptide #3 having G-CSF and TPO activity.

KW Fusion protein; human granulocyte colony stimulating factor; hG-CSF;
KW thrombopoietin; TPO; spacer peptide; blood platelet production;
KW leukocyte production; anaemia; ds.

OS Homo sapiens.

key	Location/Qualifiers
sig_peptide	1..63
	/*tag= a
mat_peptide	64..1095
	/*tag= b

PN WO9634016-A1.

31-OCT-1996.

26-APR-1996: 96WO-JP01157.

PR 26-APR-1995: 95JP-0102625.

PA (KYOW) KYOWA HAKKO KOGYO KK.

PT Anazawa H. Konishi N. Shiot

PI Uchida K, Yamasaki M, Yamashita K, Yokoi H;
XX

DR WPI; 1996-49/5/3/49.
DR P-PSDB. AAW00379

XX	Fusion peptide having G-CSF and with thrombopoietin activity -
PT	optionally chemically modified with a polyalkylene glycol, used for
PT	treatment of anaemia

XX
PS
Claim 3: Page 52-54: 71pp: Japanese.

The sequences given in AAT41786-88 encode fusion proteins which are composed of a peptide having human granulocyte colony stimulating factor (hG-CSF) activity fused with a peptide having thrombopoietin (TPO) activity, opt. via a spacer peptide. Peptides derived from these by deletion, insertion or substitution of one or more amino acid residues are included within the scope of the invention. The fusion peptides stimulate blood platelet and leukocyte production and are useful in the treatment of anaemia.

Sequence 1095 BP: 194 A; 362 C; 311 G; 228 T; 0 other;

Query Match 94.3%; Score 987.4; DB 17; Length 1095;
Best Local Similarity 95.5%; Pred. No. 4,1e-227;
Matches 1046; Conservative 0; Mismatches 1; Indels 48; Gaps 1;

QY 1 ATGGAGTGAAGTGAATGCTCTCGTGGTCATGCTTCTTAACGCAAGGCTAAGCGTGG 60
Db 1 atggagtgactgaattgctctcgttgatgcttctcttaactgaaggttaacgctg 60

QY 61 TCACAGCCGGCTCCTCTGCTGCTGACCTCCGAGTCCAGTAAACTGCTGCTGACTCC 120
Db 61 tcacagccggctcctctgctgctgaccccgagtgctgctgactcc 120

QY 121 CATGCTCTTACAGCAGACTGAGCCAGTCCAGAGGTTTCAACCTTTGCTTACACTGTC 180
Db 121 catgctcttcaagcagactgagccagtgccagaggttcaaccttgcctacactgctc 180

QY 181 CTGCTGCTGCTGAGACTTACCTTGGGAGATGAAACCCAGATGGAGGAGACCAAG 240
Db 181 ctgctgctgctgagacttaccttgggagatgaaacccagatggagagaccaag 240

QY 241 GCACAGACATTTCTGGGAGCAGTGACCTTCTGCTGGAGGAGTGTGCGACGCGGGA 300
Db 241 gcacagacattctgggagcagtgaccttctgctggagggagtgatgacgagcgga 300

QY 301 CAACTGGGACCCACTTGCCTCTCATCCCTCTGGGAGCTTTCTGGACAGTCCGCTCTC 360
Db 301 caactgggacccacttgcctctcatccctctgggagcgtttctggacagtgctctc 360

QY 361 CTGCTTGGGCGCTGAGAGCCTCTTGGAAACCCAGCTTCTCCACAGGCGAGGACCA 420
Db 361 ctgcttgggcgctgagagcctcttggaaacccagcttctccacagggcagaccaca 420

QY 421 GCTCACAAGGATCCCAATGCCATCTTCTGAGTCCCAACCTGCTCCGAGGAAAGTG 480
Db 421 gctcacaaggatcccaatgccatcttctgagctcccaacctgctccgaggaagtg 480

QY 481 CGTTTCTGATGCTTAGGAGGCTCCACCCCTCTGCTGCTACGG- 522
Db 481 cgtttctgatgcttaggaggtccacccctctgctgctacgg- 522

QY 523 -----CGGGCGCCAACTATATGCGCGCTCGAGTCTA 552
Db 523 -----CGGGCGCCAACTATATGCGCGCTCGAGTCTA 552

QY 553 CCACAGAGTCTCTTTTAAAGCTTAGAGCAAGTGGAGAACTCCAGGCGCATGCGCA 612
Db 553 ccacagagtctctttttaaagcttagagcaagtggagaaactccaggcgcatgcgca 612

QY 601 CCACAGCTCTCTTAAAGCTTAGAGCAAGTGGAGAACTCCAGGCGCATGCGCA 660
Db 601 ccacagctctctttaaagcttagagcaagtggagaaactccaggcgcatgcgca 660

QY 613 CGGCTCCAGGAGCTGTGCTGCCACTACAGCTGTGCCACCCGAGGAGCTGTGCTG 672
Db 613 cggctccaggagctgtgctgccactacagctgtgccacccgaggagctgtgctg 672

QY 661 GCGCTCCAGGAGCTGTGCTGCCACTACAGCTGTGCCACCCGAGGAGCTGTGCTG 720
Db 661 gcgctccaggagctgtgctgccactacagctgtgccacccgaggagctgtgctg 720

QY 673 CTCGGACACTCTGCGGATCCCTGCGGCTCCCTGAGGAGTCCGCGGAGGCGGCTG 732
Db 673 ctcggacactctgcggaatccctgcggtccctgagcagtgccccagcgagcgctg 732

QY 721 CTCGGACACTCTGCGGATCCCTGCGGCTCCCTGAGGAGTCCGCGGAGGCGGCTG 780
Db 721 ctcggacactctgcggaatccctgcggtccctgagcagtgccccagcgagcgctg 780

QY 733 CAGCTGGAGGCTGTGAGGCAACTCCATAGCGGCTTTCTCTACCGAGGCTCTCTG 792
Db 733 cagctggaggtgtgagggcaactccatagcggtttctctacccgaggtctctg 792

QY 781 CAGCTGGAGGCTGTGAGGCAACTCCATAGCGGCTTTCTCTACCGAGGCTCTCTG 840
Db 781 cagctggaggtgtgagggcaactccatagcggtttctctacccgaggtctctg 840

QY 793 CAGGCGCTGGAGGATCTCCCGGAGTGGGTCCACCTTGGACACACTGACGCTGGAC 852
Db 793 caggcgctggaggatctcccgagtggtgggtccaccttgacacactgacgctggac 852

QY 841 CAGGCGCTGGAGGATCTCCCGGAGTGGGTCCACCTTGGACACACTGACGCTGGAC 900
Db 841 caggcgctggaggatctcccgagtggtgggtccaccttgacacactgacgctggac 900

QY 853 GTCCGCGACTTGTCCACACATCTGGCAGCAGATGGAGAACTGGGAATGCCCTGCC 912
Db 853 gtccgcgacttgtccacacatctggcagcagatggagaaactgggaatgccctgcc 912

QY 901 GTCGCGAGCTTGTGCAACCACTCTGCGAGCAGTGGAGAACTGGGAATGCCCTGCC 960
Db 901 gtcgcgagcttgtgcaaccactctgagcagcagatggagaaactgggaatgccctgcc 960

QY 913 CTCACGCCACCCAGGCTGCCATGCGGCTTCTGCTCTGCTTTCACGCGGCGGACGA 972
Db 913 ctcacgccacccaggctgccatgcggtctctgctcttccacgcttcacgcttcacg 972

QY 961 CTCGCGAGCTTGTGCAACCACTCTGCGAGCAGTGGAGAACTGGGAATGCCCTGCC 1020
Db 961 ctcgcgagcttgtgcaaccactctgagcagcagatggagaaactgggaatgccctgcc 1020

QY 973 GGGGTCTAGTGGCTTCCATCTGCAGAGCTTCTGCTGGAGGCTGCTACGC 1032
Db 973 ggggtctagtggcttccatctgcagagcttctgctggaggtgctacgc 1032

Db 1021 ggggtctagtggctcccatctgcagagcttctctgaggtgctgacgcgttctacgc 1080

QY 1033 CACCTTGCCAGCCC 1047
Db 1081 cacttgccagccc 1095

RESULT 4
AAQ99552
ID AAQ99552 standard; cDNA to mRNA; 861 BP.
XX
AC AAQ99552;
XX
DT 29-APR-1996 (first entry)
XX
DE Thrombopoietin coding sequence.
XX
KW Thrombopoietin; TPO; humoral factor; platelet; antibody; therapy; AIDS;
KW Thrombocytopenia; hypoplastic anaemia; thrombotic thrombocytopenia;
KW disseminated intravascular coagulation syndrome; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 25..783
FT /tag= a
FT /product= thrombopoietin
FT sig_peptide 25..87
FT /tag= b
FT mat_peptide 88..783
FT /tag= c
XX
PN WO9521919-A2.
XX
PD 17-AUG-1995.
XX
PF 14-FEB-1995; 95WO-JP00208.
XX
PR 14-FEB-1994; 94JP-0039090.
PR 14-MAR-1994; 94US-0212164.
PR 25-MAR-1994; 94JP-0079842.
PR 01-APR-1994; 94US-0221020.
PR 01-JUN-1994; 94JP-0155126.
PR 15-JUN-1994; 94JP-0167328.
PR 20-JUL-1994; 94US-0278083.
PR 17-AUG-1994; 94JP-0193169.
PR 18-AUG-1994; 94JP-0193169.
PR 11-OCT-1994; 94US-0320300.
PR 01-NOV-1994; 94JP-0304167.
PR 01-DEC-1994; 94JP-0298669.
PR 22-DEC-1994; 94US-0361811.
PR 28-DEC-1994; 94JP-0341200.
PR 31-JAN-1995; 95US-0381478.
PR 17-AUG-1994; 94JP-0227159.
XX
(KIRI) KIRIN BREWERY KK.
XX
PI Akahori H, Iwamatsu A, Kato T, Kuroki R, Miyazaki H;
PI Muto T, Ohgami K, Shimizu T;
XX
WPI; 1995-293120/38.
DR P-PSDB; AAR81378.
XX
Thrombopoietin polypeptide which specifically stimulates or
increases platelet production - useful in treatment of
thrombocytopenia, also related DNA and vectors
Claim 19; Page 337-338; 383pp; English.
XX
This sequence represents the coding sequence for a thrombopoietin (TPO)
protein. TPO is a humoral factor capable of promoting platelet
production. The encoded protein can be purified by using a gel affinity

CC column where Mpl has been coupled to a resin. AAQ99551 and AAQ99553
CC also encode TPO proteins. The DNA sequences are inserted into vectors
CC which are used to transform prokaryotic and eukaryotic host cells. Using
CC the proteins encoded by these sequences, and derivatives of them,
CC antibodies specifically immunoreactive with a TPO protein can be created.
CC The antibodies, DNA sequences and vectors are used to isolate the
CC protein sequences. The TPO proteins can then be used in the treatment
CC of platelet disorders. These include thrombocytopenia, hypoplastic
CC anaemia, AIDS, disseminated intravascular coagulation syndrome and
CC thrombotic thrombocytopenia.

XX
SQ Sequence 861 BP; 243 A; 257 C; 197 G; 164 T; 0 other;

Query Match 50.3%; Score 526.6; DB 16; Length 861;
Best Local Similarity 99.2%; Pred. No. 7.2e-117;
Matches 529; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGGAGCTGACTGAATGTCTCTCGTGGTCAATGCTTCTTCTCACTGCAAGGCTAAAGCGTG 60
DB 25 atggagctgactgaattgctctcggtggtcattctctcctaactgcaaggctaaagctg 84

QY 61 TCAGAGCCGCTCTCTGCTGTGACCTCCGAGTCTCAGTAACTGCTTCTGACTCC 120
DB 85 tcagagccgctctctctgctgtgacctccgagctcagtaaaactgctctgactcc 144

QY 121 CATGTCTTCCACAGCAGTACGAGTCCAGAGTTCACCTTTGCTACACCTGTC 180
DB 145 catgtcttccacagcagctgagccagtgccagaggttcacaccttggctacacctgctc 204

QY 181 CTGCTGCTGCTGCTGACTTTAGCTTGGGAGATGGAATGGAACCCAGATGGAGGACCAAG 240
DB 205 ctgctgctgctgactttgacttggagatggagaaaccagatggagagaccaag 264

QY 241 GCACAGACATTTGGAGCAGTACCTTCTGCTGGAGGAGTATGCGACGACGGGGA 300
DB 265 gcacagacattctggagcagtgacctctctgctggagagtgatggcagcacgggga 324

QY 301 CAACTGGGACCCACTTGCCTCTCATCTCCCTGSGGAGCTTTCGACAGGTCCTGTC 360
DB 325 caactgggacccacttgcctctcatcctctctggcaggttcttgagacaggtccgctc 384

QY 361 CTCCTTTGGGCGCTGACAGACCTTCTTGGAAACCCAGCTTCTCCACAGGCGACACCA 420
DB 385 ctcttggggccctgcagagctctcttgaaacccagcttctccacagggcagaccaca 444

QY 421 GCTCAAGAGATCCCAATGCCATCTTCTGAGCTTCCACACCTGCTCCGAGGAAGGTG 480
DB 445 gctcaagagatcccaatgccatctctctgagcttccacacactgctccgaggaaggtg 504

QY 481 CGTTTCTGATGCTGTAGGAGGTCCACCTCTGCGTAGCGGGCGGCCAAC 533
DB 505 cgttctctgatgctgttaggaggggtccacctctgctgagcgggcccccacc 557

RESULT 5
ID AAT03941
XX AAT03941 standard; cDNA; 1062 BP.
XX AAT03941;
AC AAT03941;
DT 28-APR-1996 (first entry)
XX Human thrombopoietin coding sequence.
DE
XX Thrombopoietin; erythropoiesis stimulator; treatment;
KW thrombocytopenia; anaemia; ds.
XX
OS Homo sapiens.
XX

Key Location/Qualifiers
FH 1..1062
FT CDS /*tag= a

XX WO9521626-A1.
PN
XX 17-AUG-1995.
PD
XX
XX 09-FEB-1995; 95WO-US01829.
XX
XX 01-DEC-1994; 94US-0347748.
PR 14-FEB-1994; 94US-0196025.
PR 25-FEB-1994; 94US-0203197.
PR 21-MAR-1994; 94US-0215203.
PR 01-JUN-1994; 94US-0252491.
PR 09-AUG-1994; 94US-0288417.
PR 07-NOV-1994; 94US-0335566.
XX (UNIW) UNIV WASHINGTON.
XX
XX Kaushansky K;
PI
XX
XX WPI; 1995-292944/38.
DR P-PSDB; AAR82682.
XX
XX Stimulation of erythropoiesis using thrombopoietin and opt.
PT erythropoietin - for the treatment of thrombocytopenia and anaemia.
XX
XX Disclosure; Page 38-40; 66pp; English.
XX
XX This sequence encodes human thrombopoietin, which stimulates
CC erythropoiesis to produce an increase in proliferation of
CC differentiation of erythroid cells or to increase reticulocyte
CC counts at least 2-fold over baseline reticulocyte counts and,
CC optionally, platelet levels to at least 20000/cu mm. The protein
CC can be used in a composition, optionally with erythropoietin, for
CC use in the treatment of thrombocytopenia and anaemia, such as
CC that caused by destruction of haematopoietic cells in bone marrow,
CC in the treatment of cancer with chemotherapy and radiation, and in
CC pathological conditions such as myelodysplasia, AIDS, aplastic
CC anaemia, autoimmune disease or inflammatory disease.
XX
XX Sequence 1062 BP; 223 A; 375 C; 236 G; 228 T; 0 other;

Query Match 50.3%; Score 526.6; DB 16; Length 1062;
Best Local Similarity 99.2%; Pred. No. 7.5e-117;
Matches 529; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGGAGCTGACTGAATGTCTCTCGTGGTCAATGCTTCTTCTCACTGCAAGGCTAAAGCGTG 60
DB 1 atggagctgactgaattgctctcggtggtcattctctcctaactgcaaggctaaagctg 60

QY 61 TCAGAGCCGCTCTCTGCTGTGACCTCCAGTCTCAGTAACTGCTTCTGACTCC 120
DB 61 tcagagcccgctctctgctgtgacctccgagctcagtaaaactgctctgactcc 120

QY 121 CATGTCTTTCACAGCAGTACGAGTCCAGAGTTCACCTTTGCTTACACCTGTC 180
DB 121 catgtcttccacagcagactgagcaggtgcccagaggttcaccttggctacacctgctc 180

QY 181 CTGCTGCTGCTGCTGGACTTTAGCTTGGGAGATGGAATGGAACCCAGATGGAGGACCAAG 240
DB 181 ctgctgctgctgctggactttagcttgggagaaatggaaccacagatggaggagaccaag 240

QY 241 GCACAGACATTTCTGGAGCAGTGACCTTCTGCTGAGGAGTATGCGACGACGGGGA 300
DB 241 gcacagacattctctggagcagtgacctctctgctggagaggtatggcagcacgggga 300

QY 301 CAACTGGGACCCACTTGCCTCTCATCTCCCTGSGGAGCTTTCGACAGGTCCTGTC 360
DB 301 caactgggacccacttgcctctcatcctctctgggagagcttcttgagacaggtccgctc 360

QY 361 CTCCTTTGGGCGCTGACAGACCTTCTTGGAAACCCAGCTTCTCCACAGGCGACACCA 420
DB 361 ctcttggggccctgcagagctctcttggaaacccagcttctccacagggcagaccaca 420

QY 421 GCTCACAGGATCCCAATGCCATCTTCTCTGAGTTCACACACCTGCTCCGAGGAAAGGTG 480
 Db 421 gctcacagagatcccaatgccatcttctctgagcttccacacactgctccgaggaaggtg 480
 QY 481 CGTTTCTGATGTTGTAGGAGGTTCCACCCTCTGCGTACGGCGGGCGCCCAAC 533
 Db 481 cgtttctgatgctttaggaggttccaccctctgctcagcgggcccccacc 533

RESULT 6

AAT04050
 ID AAT04050 standard; cDNA; 1062 BP.

AC AAT04050;

XX 28-MAR-1996 (first entry)

XX Sequence encoding haematopoietic protein.

XX Haematopoiesis; protein; bone marrow; stem cells; precursor;
 KW platelet; therapy; ds.

XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS 1..1062

XX /tag= a

XX /product= Haematopoietic protein.

XX W09521920-Al.

XX PN 17-AUG-1995.

XX 05-AUG-1994; 94WO-US08806.

XX 01-JUN-1994; 94US-0252491.

XX 14-FEB-1994; 94US-0196025.

XX 25-FEB-1994; 94US-0203197.

XX 21-MAR-1994; 94US-0215203.

XX (UNIW) UNIV WASHINGTON.

XX (ZYMO) ZYMOGENETICS INC.

XX Burkhead SK, Foster DC, Hagen FS, Holly RD, Kaushansky K;

XX Kuijper JL, Lofton-day C, Lok S, Oort PJ;

XX WPI; 1995-293121/38.

XX P-PSDB; AAR79906.

XX Haematopoietic proteins and polypeptide(s) - useful for in vivo and

XX ex vivo therapy

XX Claim 20; Page 106-108; 137pp; English.

XX Haematopoietic proteins and polypeptides are useful for stimulating

XX platelet production in a mammal when given in a therapeutically

XX effective amount. They are also useful for stimulating bone marrow

XX cell proliferation where the bone marrow cells are megakaryocytes or

XX their precursors.

XX Sequence 1062 BP; 223 A; 375 C; 236 G; 228 T; 0 other;

XX Query Match 50.3%; Score 526.6; DB 16; Length 1062;

XX Best Local Similarity 99.2%; Pred. No. 7.5e-117;

XX Matches 529; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGGAGCTGACTGAATTGCTCTGCTGCTCATGCTTCTCTCACTGCAAGGCTAACGCTG 60

Db 1 atggagctgactgaattgctctgctgctcatgcttctctcaactgcaagctgaacgctg 60

QY 61 TCCAGCCCGGCTCCCTCCCTGCTGTGACCTCGAGTCTCAGTAACACTGCTTCGTGACTCC 120

Db 61 tccagcccggtccctccctgcttgtgacctccgagtcctcagtaaacctgtctctgactcc 120
 QY 121 CATGTCTTTCACAGCAGACTGAGCCAGTGCCAGAGGTTTACCCCTTTCCTACACCTGTC 180
 Db 121 catgtcttctcacagcagactgagccagtgcccagaggttcaccccttgcctacacctgtc 180
 QY 181 CTGCTGCCCTGTGTGACTTTAGCTTGGGAGAAATGGAACCCAGATGGAGGAGACCAAG 240
 Db 181 ctgctgccctgtgtgacttttagcttggagaatggaaacccagatggaggagaccaag 240
 QY 241 GCACAGGACATTTCTGGGACAGTACCCCTTCTGCTGGAGGAGTGTGTCGAGCAGCGGGA 300
 Db 241 gcacaggacattctggagcagtacccttctgctggaggagtgatggcagcacggga 300
 QY 301 CAACTGGGACCCACTTGCTCTCATCCCTCTGCTGGGCGACGCTTCTCGACAGGTCCTCTC 360
 Db 301 caactgggaccacttgctctcatccctctcctcctcctgggagcgttctctggaggtcccgctc 360
 QY 361 CTCCTTGGGGCCCTGCAGAGCCTCTTGAACCCAGCTTCTCCACAGGCGAGGACACA 420
 Db 361 ctcttggggccctgcagagcctccttggaaacccagcttctccacagggcaggaccaca 420
 QY 421 GCTCACAAAGGATCCCAATGCCATCTTCTGAGCTTCCAAACACTGCTCCGAGGAAAGGTG 480
 Db 421 gctcacaaaggatcccaatgccatcttctgagcttccaaacactgctccgaggaaaggtg 480
 QY 481 CGTTTCTCTGATGTTGTAGGAGGTTCCACCCTCTGCGTACGGCGGGCGCCCAAC 533
 Db 481 cgtttctgatgctttaggaggttccaccctctcctcctcctcctcctcctcctcctcctcctc 533

RESULT 7

AAT34852
 ID AAT34852 standard; cDNA; 1062 BP.

AC AAT34852;

XX 03-DEC-1996 (first entry)

XX Human thrombopoietin cDNA.

XX Thrombopoietin; TPO; thrombocytopenia; platelet; therapy; ds.

XX Homo sapiens.

XX W09617062-Al.

XX 06-JUN-1996.

XX 15-NOV-1995; 95WO-US14929.

XX 30-NOV-1994; 94US-0346999.

XX (ZYMO) ZYMOGENETICS INC.

XX Buddle MM, Downey W, Forstrom JW, Foster DC, Lok S;

XX Osborn SG;

XX WPI; 1996-286830/29.

XX P-PSDB; AAR98947.

XX New low mol. wt. forms of thrombopoietin and related DNA - vectors

XX and transformed cells, useful for stimulating platelet prodn. esp.

XX for treatment of thrombocytopenia

XX Claim 19; Page 57-59; 84pp; English.

XX A cDNA clone (AAT34852) codes for human thrombopoietin (AAR98947), a

XX protein that stimulates megakaryopoiesis and thrombocytopoiesis.

XX The cDNA can be incorporated into an expression vector and utilised

XX in the prodn. of low mol.wt. bioactive human thrombopoietin in

XX eukaryotic (partic. yeast or mammalian) host cells. The recombinant

XX

CC This restoration may be further improved by administering TPO to the
XX recipient after transplant. The present sequence encodes human TPO.
SQ Sequence 1062 BP; 223 A; 375 C; 236 G; 228 T; 0 other;

Query Match 50.3%; Score 526.6; DB 18; Length 1062;
Best Local Similarity 99.2%; Pred. No. 7.5e-117;
Matches 529; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGAGCTGACTGAATGCTCTCGTGTGATGCTTCTCTTAAGCTGAAGCTTAACGCTG 60
Db 1 atggagctgactgaatgctctcgtgtgatgcttctcttaagctgaagcttaacgctg 60
QY 61 TCCAGCGCGCTCTCTGCTGTGAGCTCCGAGTCCCTCAGTAAACTGCTTCGTGACTCC 120
Db 61 tccagcgcgcctctctgctgtgagctccgagctccctcagtaaacctgcttcgtgactcc 120
QY 121 CATGCTCTTACAGCAGACTGAGCCAGTGGCCAGAGGTTTACCCCTTTGCCCTACACCTGTC 180
Db 121 catgctcttacagcagactgagccagtgccagaggttcaacctttgcttaacctgtc 180
QY 181 CTGCTGCTGTGAGCTTTAGCTTGGGAGAAATGGAAACCCAGATGGAGGAGACCAAG 240
Db 181 ctgctgctgtgagcttttagcttgggagaaatggaaacccagatggagagaccag 240
QY 241 GCACAGGACATTCCTGGGAGAGTACACCTTCTGCTGGAGGAGTGGCAGCAGCGGGA 300
Db 241 gcacagacatctcctgggagagtgacctctctgctgagggagtgatggcagcacgggga 300
QY 301 CAATGGGACCCACTTGCCTCTCATCCCTCTCTGGGGCAGGCTTCTGGACAGGTCCTGTC 360
Db 301 caatgggacccacttgcctctcatccctctctggggcaggttctggacaggtccgtc 360
QY 361 CTCTTTGGGCGCTGACAGCTCTTGGAAACCCAGCTTCTTCCACAGGGCAGGACCA 420
Db 361 ctctttgggcgctgacagctcttggaaacccagcttcttccacagggcaggaccaca 420
QY 421 GCTCACAGGATCCCAATGCATCTCTGAGCTTCCAAACACCTGCTCCGAGGAAAGTG 480
Db 421 gctcacaggatcccaatgcctctctgagcttccaaacacctgctccgaggaaggtg 480
QY 481 CGTTTCTGATGCTTTAGGAGGTCACCCCTCTGCTAGCGGGCGGCCAAC 533
Db 481 cgtttctgatgcttttaggaggtccacccctctgctcaggcgggcccccacc 533

RESULT 11

AAV21696
AAV21696 standard; cDNA; 1062 BP.

AAV21696;

07-AUG-1998 (first entry)

Human thrombopoietin encoding cDNA.

Thrombopoietin; TPO; expression vector; platelet number; cytokine;
human; bone marrow; proliferation; treatment; cancer; aplastic anaemia;
myelodisplastic syndrome; chemotherapy; cytopenia; thrombocytopenia;
haematologic disorder; leukaemia; lymphoma; ss.

Homo sapiens.

Key	Location/Qualifiers
PH CDS	1..1062
FT	/*tag= a
FT	/product= "Thrombopoietin polypeptide"
FT	1..63
FT	/*tag= b
FT	64..1059
FT	/*tag= c

PN WO9806849-A1.
XX 19-FEB-1998.
XX 30-JUL-1997; 97WO-US13543.
XX 13-AUG-1996; 96US-0696447.
XX (ZYMO) ZYMOGENETICS INC.
XX Irani M, Morrison-nelson GR;
XX WPI; 1998-159541/14.
XX P-PSDB; AAW53124.
XX Thrombopoietin protein expression vector - used for increasing
XX platelet number in a mammal
XX Disclosure; Pages 29-32; 56pp; English.

CC This cDNA encodes a thrombopoietin (TPO) polypeptide. This can be used
CC in the construction of a new expression vector replicable in a
CC eukaryotic host cell encoding TPO polypeptides. The vector comprises a
CC transcription promoter, a first DNA segment encoding a secretory leader,
CC a second segment encoding a TPO polypeptide and a transcription
CC terminator which are all operably linked. The second TPO segment consists
CC of C-X-B, where C is a human TPO cytokine domain, X is a peptide bond or
CC a linker consisting of one or two amino acid residues, such that X along
CC in combination with C or B does not provide a dibasic amino acid pair
CC and B is a polypeptide that can be selected from AAW53125 to AAW53126.
CC The secretory leader is a S. cerevisiae alpha-factor secretory leader. A
CC cultured yeast cell containing such an expression vector can be used to
CC produce the TPO polypeptide. The TPO polypeptide can be used in a method
CC for increasing platelet number in a mammal. It can be used to increase
CC proliferation of bone marrow cells for treatment of cytopenia, including
CC those induced by aplastic anaemia, myelodisplastic syndromes,
CC chemotherapy or congenital cytopenias. It can also be used to treat
CC thrombocytopenia, haematologic disorders, such as leukaemia and lymphoma
CC or metastatic cancers involving bone marrow.

SQ Sequence 1062 BP; 223 A; 375 C; 236 G; 228 T; 0 other;

Query Match 50.3%; Score 526.6; DB 19; Length 1062;
Best Local Similarity 99.2%; Pred. No. 7.5e-117;
Matches 529; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

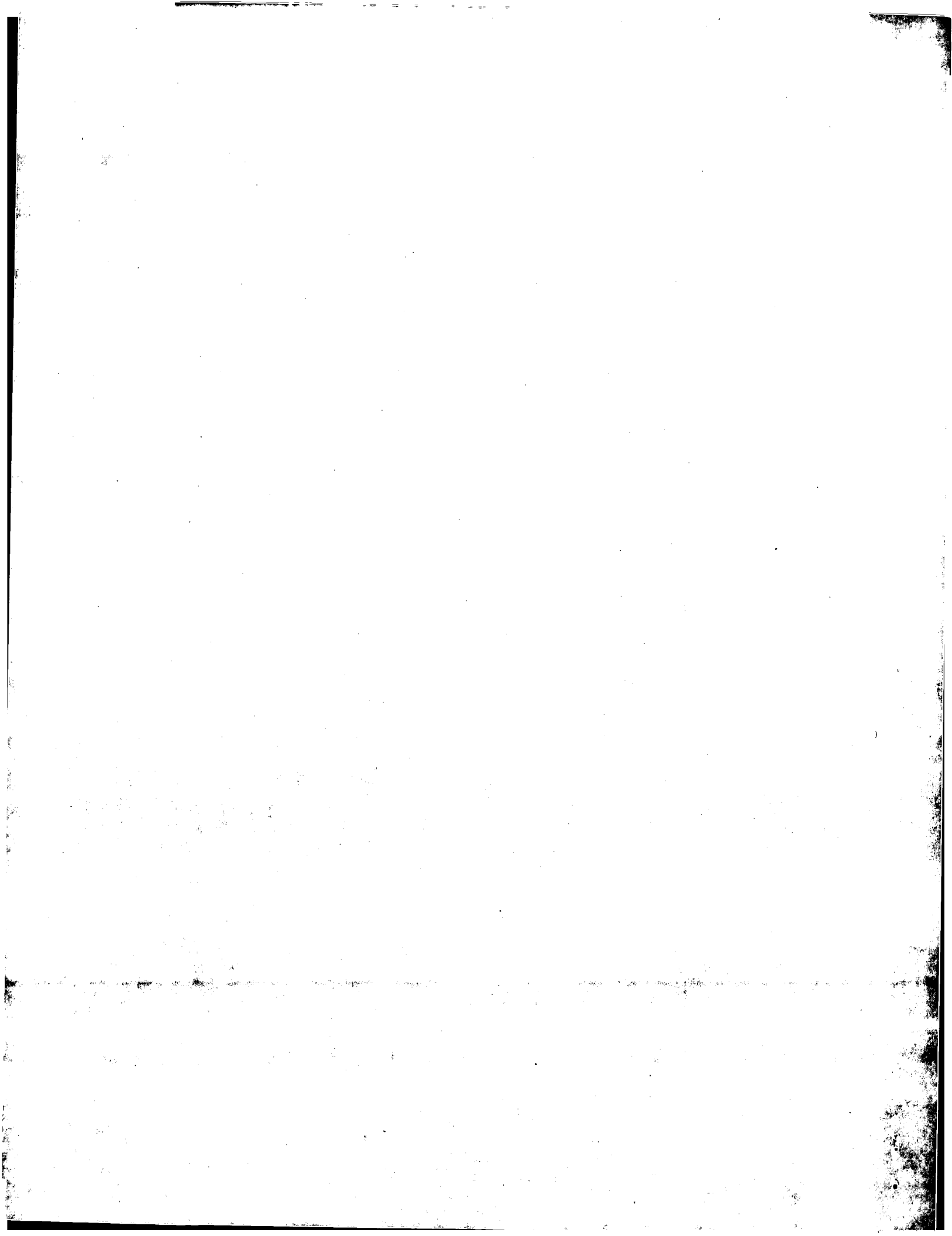
QY 1 ATGAGCTGACTGAATGCTCTCGTGTGATGCTTCTCTTAAGCTGAAGCTTAACGCTG 60
Db 1 atggagctgactgaatgctctcgtgtgatgcttctcttaagctgaagcttaacgctg 60
QY 61 TCCAGCGCGCTCTCTGCTGTGAGCTCCGAGTCCCTCAGTAAACTGCTTCGTGACTCC 120
Db 61 tccagcgcgcctctctgctgtgagctccgagctccctcagtaaacctgcttcgtgactcc 120
QY 121 CATGCTCTTACAGCAGACTGAGCCAGTGGCCAGAGGTTTACCCCTTTGCCCTACACCTGTC 180
Db 121 catgctcttacagcagactgagccagtgccagaggttcaacctttgcttaacctgtc 180
QY 181 CTGCTGCTGCTGTGAGCTTTAGCTTGGGAGAAATGGAAACCCAGATGGAGGAGACCAAG 240
Db 181 ctgctgctgctgtgagcttttagcttgggagaaatggaaacccagatggagagaccag 240
QY 241 GCACAGGACATTCCTGGGAGGAGTACACCTTCTGCTGGAGGAGTGGCAGCAGCGGGA 300
Db 241 gcacagacatctcctgggagagtgacctctctgctgagggagtgatggcagcacgggga 300
QY 301 CAACTGGGACCCACTTGCCTCTCATCCCTCTGCTGGGCGAGCTTCTTGACAGGTCCTGTC 360
Db 301 caactgggacccacttgcctctcatccctctctgctggggcaggttcttgagacaggtccgtc 360
QY 361 CTCTTTGGGCGCTGACAGCTCTTGGAAACCCAGCTTCTTCCACAGGGCAGGACCA 420
Db 361 ctctttgggcgctgacagctcttggaaacccagcttcttccacagggcaggaccaca 420

Db 361 ctcttggggccctgcagagcctcttggaaaccagcttctccacagggcaggaccaca 420
 QY 421 GCTCACAGGATCCCAATGCATCTTCTCAGCTTCCAAACACTGCTCCGAGGAAAGGTG 480
 Db 421 gctcacagagatcccaatgccatcttccgagcttccacacctgctcggaggaaggcg 480
 QY 481 CGTTTCTGATGCTTGTAGGAGGTCACCCCTCTGCGTACGGCGGCCCAAC 533
 Db 481 cgttctctgatctgttaggaggggtccacctctgctcagcgggcccccacc 533

RESULT 12
 AAA51991
 ID AAA51991 standard; cDNA; 1062 BP.
 XX
 AC AAA51991;
 XX
 DT 04-DEC-2000 (first entry)
 XX
 XX Human thrombopoietin cDNA.
 XX
 KW Thrombopoietin; TPO; erythropoietin; EPO; erythropoiesis; stimulate;
 KW anaemia; radiation; chemotherapy; red blood cell production;
 KW haematopoietic cell; cytopenia; ex vivo expansion; anti-anaemic;
 KW immunomodulator; ds.
 XX
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FH 1..1062
 FT CDS
 FT /*tag= a
 FT /product= Thrombopoietin
 XX
 XX US6099830-A.
 XX
 XX 08-AUG-2000.
 XX
 XX 12-SEP-1997; 97US-0999927.
 XX
 XX 09-AUG-1994; 94US-0288417.
 XX 07-NOV-1994; 94US-0335566.
 XX
 XX (UNIW) UNIV WASHINGTON.
 XX (ZYMO) ZYMOGENETICS INC.
 XX
 XX Kaushansky K;
 XX
 XX WPI; 2000-557618/51.
 XX P-PSDB; AAY97136.
 XX
 XX Stimulating erythropoiesis e.g. for treating anemias or cytopenia, or
 XX for ex vivo expansion of marrow or peripheral blood progenitor cell,
 XX comprises administering thrombopoietin alone or in combination with
 XX erythropoietin
 XX
 XX Disclosure; Column 19-22; 23pp; English.
 XX
 XX Erythropoiesis may be stimulated in anaemic mammals by administration of
 XX thrombopoietin (TPO), and optionally erythropoietin (EPO). The anaemic
 XX mammal to be treated will have shown a haematocrit level less than 33
 XX percent of total blood volume, a reticulocyte count of less than
 XX 0.8 percent and may have been treated with radiation or chemotherapy. The
 XX improved red blood cell level is due to an increase in red blood cell
 XX production rather than a decrease in red blood cell destruction. TPO
 XX proteins may be used therapeutically to increase proliferation of
 XX haematopoietic cells in the bone marrow, such as in the treatment of
 XX cytopenia or anaemia, especially associated with bone marrow failure,
 XX disorders associated with low red blood cell production, particularly
 XX when accompanied by thrombocytopenia and for increasing the level of
 XX circulating erythrocytes and erythrocyte precursor cells. TPO and EPO may
 XX be used for ex vivo expansion of marrow or peripheral blood progenitor
 XX cells.

SQ Sequence 1062 BP; 223 A; 375 C; 236 G; 228 T; 0 other;
 Query Match 50.3%; Score 526.6; DB 21; Length 1062;
 Best Local Similarity 99.2%; Pred. No. 7.5e-117; Indels 0; Gaps 0;
 Matches 529; Conservative 0; Mismatches 4;
 QY 1 ATGGAGCTGACTGAATTGCTCGTGGTCACTGCTTCTTCTTAAGTCAAGGCTAAACGCTG 60
 Db 1 atggagctgactgaattgctcctcggtcatgcttctcctaactgcaaggtcaacgctg 60
 QY 61 TCCAGCCCGGCTCCTCCTGCTGTGACCTCCGAGTCTCTAGTAACTGCTTCTGTGACTCC 120
 Db 61 tcagcccggtcctcctctgctgtgacctccgagtcctcagtaactgcttctgactcc 120
 QY 121 CATGTCTTTCACAGCAGACTGAGCCAGTCCCGAGAGTTCACCCCTTTCCTACACCTGTC 180
 Db 121 catgtcttccacagcagactgagccagtgccccagaggttcaccccttgcctacacctgc 180
 QY 181 CTCCTGCCCTGCTGCTGACTTTAGCTTTGGGAGATGAAACCCAGATGGAGGAGACCAAG 240
 Db 181 ctctgcctgctgctgagactttagcttggagaatgaaaacccagatggagagaccaag 240
 QY 241 GCACAGGACATTTCTGGGACAGTACCCTTCTGCTGGAGGAGTGTGCGCAGCACGGGGA 300
 Db 241 gcacaggacattctggagcagtgaccttctgctgaggagtgatggcagcacgggga 300
 QY 301 CAACTGGGACCCACTTGCCTCTCATCTCCCTCGGCGAGCTTTCTGGACAGGTCCGTC 360
 Db 301 caactgggacccacttgcctctcatctccctcctcctgggagcgttcttggacaggtcc 360
 QY 361 CTCCTTGGGGCCCTGCAGAGCCTCTTGGAAACCCAGCTTCTCCACAGGCGAGSACCACA 420
 Db 361 ctcttggggccctgcagagcctcttggaaacccagcttctccacagggcaggaccaca 420
 QY 421 GCTCACAGGATCCCAATGCCATCTTCTGAGCTTCCACACCTGCTCCGAGGAAAGGTG 480
 Db 421 gctcacagagatcccaatgccatcttctgagcttccacacctgtctcggaggaagggtg 480
 QY 481 CGTTTCTGATGCTTGTAGGAGGTCACCCCTCTGCGTACGGCGGCCCAAC 533
 Db 481 cgttctctgatctgttaggaggggtccacctctgctcagcgggcccccacc 533

RESULT 13
 AAT47958
 ID AAT47958 standard; cDNA to mRNA; 1086 BP.
 XX
 AC AAT47958;
 XX
 DT 03-APR-1997 (first entry)
 XX
 DE Human thrombopoietin (1-332) coding sequence.
 XX
 KW TPO; recombinant; thrombopoietin; storage; reduced adsorption;
 KW container wall; additive; ds.
 XX
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FH 25..1086
 FT CDS
 FT /*tag= a
 FT /note= "immature protein"
 FT sig_peptide 25..87
 FT /*tag= b
 FT mat_peptide 88..1083
 FT /*tag= c
 FT /note= "mature protein residues 1-332"
 XX
 XX W09628181-A1.
 XX
 XX 19-SEP-1996.
 XX



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 5, 2002, 08:51:36 ; Search time 1909.33 seconds
(without alignments)
5892.554 Million cell updates/sec

Title: US-09-680-514-4
Perfect score: 1047
Sequence: 1 ATGGAGCTGACTGAATTGCT.....TACGCCACTTGCCCGAGCCC 1047

Scoring table: IDENTITY_NUC
Gap 10.0 , Gapext 1.0
Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estfun:*
2: em_esthum:*
3: em_estin:*
4: em_estom:*
5: em_estpl:*
6: em_estba:*
7: em_estro:*
8: em_estov:*
9: em_htc:*
10: gb_est1:*
11: gb_est2:*
12: gb_htc:*
13: gb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pin:*
18: em_gss_pro:*
19: em_gss_rod:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	320.4	30.6	973	11	BI411128
2	312.8	29.9	598	11	BG548320
3	301.6	28.8	900	11	BI328788
4	300.6	28.7	507	10	BE485194
5	300.6	28.7	633	10	AI314551
6	299.6	28.6	609	10	AA538257
7	285.2	27.2	878	11	BF789308
8	230	22.0	464	10	BE754587
9	135.8	13.0	327	11	BF848766
10	135.2	12.9	449	13	AZ025993
11	109.4	10.4	419	13	AZ352219
12	86.4	8.3	708	13	AZ792988
					2M0046105

13	77.8	7.4	752	13	AQ888740
14	65.6	6.3	589	10	BE377769
15	54.8	5.2	925	13	CNS0091P
16	50.2	4.8	513	10	BE480590
c 17	49	4.7	925	13	CNS0091P
18	48.8	4.7	617	10	AI981598
c 19	43.2	4.1	461	10	BE598945
c 20	43.2	4.1	559	11	BG556176
c 21	42.6	4.1	452	10	AA155632
c 22	42.6	4.1	899	13	CNS01VYT
c 23	42.4	4.0	1203	13	CNS015Y4
c 24	42.2	4.0	598	10	AA115932
c 25	41.4	4.0	421	10	AA058743
c 26	41.4	4.0	478	11	BI345642
27	41.2	3.9	1201	13	CNS0141B
c 28	40.8	3.9	545	11	BF868325
c 29	40.6	3.9	447	11	BG604563
c 30	40.6	3.9	845	10	AL572931
c 31	40.6	3.9	932	13	CNS0072Q
c 32	40.6	3.9	1003	11	BI410408
c 33	40.6	3.9	1101	13	CNS05709
c 34	40.4	3.9	458	10	AW273202
c 35	40.4	3.9	485	10	BE463718
c 36	40.2	3.8	864	13	CNS022LE
c 37	39.8	3.8	387	10	AW193322
c 38	39.8	3.8	427	10	BE048584
c 39	39.8	3.8	432	10	AW175624
c 40	39.8	3.8	441	10	AI818468
c 41	39.8	3.8	444	10	AW072844
c 42	39.8	3.8	450	10	AW273147
c 43	39.8	3.8	453	10	AW337946
c 44	39.8	3.8	492	13	AQ686347
c 45	39.8	3.8	505	10	AI952164

ALIGNMENTS

RESULT 1

BI411128
LOCUS 602962472F1 NCI_CGAP_Lu33 Mus musculus cdna clone IMAGE:5118185 5', mRNA EST
DEFINITION mRNA sequence.
ACCESSION BI411128.1 GI:15172051
VERSION BI411128
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 973)
NIH-MGC http://mgc.nci.nih.gov/
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgabs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11287 row: o column: 18
High quality sequence start: 32
High quality sequence stop: 912.
Location/Qualifiers
1..973
/organism="Mus musculus"
/strain="CZECH 11"
/db_xref="taxon:10090"
/clone="IMAGE:5118185"

```

/clone.lib="NCI_CGAP_Lu33"
/tissue_type="pooled lung tumors"
/lab_host="DH10B (phage-resistant)"
/note="organ: lung; Vector: pT73D-Pac (Pharmacia) with a
modified polylinker; Site_1: NotI; Site_2: EcoRI; 1st
strand cDNA was prepared from mRNA obtained from pooled
lung tumors with a Not I - oligo(dT) primer [5',
TGTTACCAATCTGAAGTCGAGCGCGGCTCTGTTTTTTTTTTTTT 3']".
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not
I and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Ronaldo. "
192 a 295 c 254 q 232 t
BASE COUNT

```

BASE COUNT	192 a	295 c	232 t	constructed	by Bento Soares and M. Fatima Bonaldo.	ORIGIN
Query Match	30.6%	Score 320.4;	DB 11;	Length 973;		
Best Local Similarity	77.1%	Pred. No. 1.8e-62;				
Matches 390;	Conservative	0;	Mismatches 116;	Indels	0;	Gaps
QY	542	CCTCAGCTCTACCACAGAGCTTCCTTTTAAAGACCTTAGACGAAGTGAGAAATCCAGG	601			
DB	91	CCCTCCCTCTGCCCGAAGCTTCCTGCTTAAGTCCCTGGAGCAAGTGAGAAATCCAGG	150			
QY	602	GCATGTGGCGACGCTCCAGGAAGGTGTGCCACCTACAAGCTGTGCCACCCGAGG	661			
DB	151	CCAGCGGCTCGGTGCTGCTGGACAGTTGTGTGCCACCTACAAGCTGTGTACCCCGAGG	210			
QY	662	AGCTGGTGTGCTCGGACACTCTCTGGGATCCCTCTGGGCTCCCTGAGCAGCTGCCCA	721			
DB	211	AGCTGGTGTGCTGGGGCACTCTCTGGGATCCCGAAGCTTCCGTGATGGCTGCTCTA	270			
QY	722	GCCAGGCCCCGTGACGTGGCAGSGTGCTTGAGGCAACTCCATAGCGGCCTTTTCTCTACC	781			
DB	271	GCCAGGCCCCGTGACGACAGACACAGTGCTTAAGCAGCTCCACAGTGGGCTCTGCCCTTACC	330			
QY	782	AGGGGCTCTCTGACGGCCCTGGAAGGATCTCCCCGAGTTGGTCCCACCTTGGACACAC	841			
DB	331	AAGGTCTCTGTGAGGCTCTATCGGATTTTCCCTCTGCCCTGGCCCCCACCCTTGGACTGC	390			
QY	842	TGCAGCTGGAGTTCGGCGACTTTGGCACCAACCATCTGGCAGCAGATGGAAGAACTGGAA	901			
DB	391	TTACAGCTGGATGTGGCAACTTTGCCACCACCACCATCTGGCAGCAGATGGAAGAACTGGGG	450			
QY	902	TGGCCCCCTGCCCTGCAGCCACCCAGGTTGCCATCCGCGCTCTGCCCTCTGCTTTCAGC	961			
DB	451	TGGCCCCCTACTGTGCAGCCACACAGAGCGCATGCCAGCTTCACTCTGCTTTCAGC	510			
QY	962	GCCGGCAGAGGGTCTAGTTGCTTCCCATCTGCAGAGCTTCCCTGGAGTGTGCTGACC	1021			
DB	511	GCCGGCAGAGGTGTCCTGGGCATTTTCGTACCTGCAGGGCTTCTCTGGAGACGCTCGCC	570			
QY	1022	GCCTTCTACGCCACCTTGCCCCAGCCC	1047			
DB	571	TTCTCTGCAACCACTTGGCCTAGACC	596			

RESULT	2	
BG548320		
LOCUS	598 bp	mrna
DEFINITION	602575289f1 NIH_MGC_77 Homo sapiens cDNA clone IMAGE:4703159 5',	EST 04-APR-2001
	mrna sequence.	
ACCESSION	BG548320	
VERSION	BG548320.1	GI:13546985
KEYWORDS	EST.	
SOURCE	human.	
ORGANISM	Homo sapiens	
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi	
REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
AUTHORS	1 (bases 1 to 598)	
	NIH-MGC http://mgc.nci.nih.gov/ .	
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)	

JOURNAL
COMMENT

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
DNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLCMI541 row: b column: 24
High quality sequence stop: 597.

FEATURES	SOURCE
1. <i>General</i>	
2. <i>Specific</i>	
3. <i>Other</i>	
4. <i>Other</i>	
5. <i>Other</i>	
6. <i>Other</i>	
7. <i>Other</i>	
8. <i>Other</i>	
9. <i>Other</i>	
10. <i>Other</i>	
11. <i>Other</i>	
12. <i>Other</i>	
13. <i>Other</i>	
14. <i>Other</i>	
15. <i>Other</i>	
16. <i>Other</i>	
17. <i>Other</i>	
18. <i>Other</i>	
19. <i>Other</i>	
20. <i>Other</i>	
21. <i>Other</i>	
22. <i>Other</i>	
23. <i>Other</i>	
24. <i>Other</i>	
25. <i>Other</i>	
26. <i>Other</i>	
27. <i>Other</i>	
28. <i>Other</i>	
29. <i>Other</i>	
30. <i>Other</i>	
31. <i>Other</i>	
32. <i>Other</i>	
33. <i>Other</i>	
34. <i>Other</i>	
35. <i>Other</i>	
36. <i>Other</i>	
37. <i>Other</i>	
38. <i>Other</i>	
39. <i>Other</i>	
40. <i>Other</i>	
41. <i>Other</i>	
42. <i>Other</i>	
43. <i>Other</i>	
44. <i>Other</i>	
45. <i>Other</i>	
46. <i>Other</i>	
47. <i>Other</i>	
48. <i>Other</i>	
49. <i>Other</i>	
50. <i>Other</i>	
51. <i>Other</i>	
52. <i>Other</i>	
53. <i>Other</i>	
54. <i>Other</i>	
55. <i>Other</i>	
56. <i>Other</i>	
57. <i>Other</i>	
58. <i>Other</i>	
59. <i>Other</i>	
60. <i>Other</i>	
61. <i>Other</i>	
62. <i>Other</i>	
63. <i>Other</i>	
64. <i>Other</i>	
65. <i>Other</i>	
66. <i>Other</i>	
67. <i>Other</i>	
68. <i>Other</i>	
69. <i>Other</i>	
70. <i>Other</i>	
71. <i>Other</i>	
72. <i>Other</i>	
73. <i>Other</i>	
74. <i>Other</i>	
75. <i>Other</i>	
76. <i>Other</i>	
77. <i>Other</i>	
78. <i>Other</i>	
79. <i>Other</i>	
80. <i>Other</i>	
81. <i>Other</i>	
82. <i>Other</i>	
83. <i>Other</i>	
84. <i>Other</i>	
85. <i>Other</i>	
86. <i>Other</i>	
87. <i>Other</i>	
88. <i>Other</i>	
89. <i>Other</i>	
90. <i>Other</i>	
91. <i>Other</i>	
92. <i>Other</i>	
93. <i>Other</i>	
94. <i>Other</i>	
95. <i>Other</i>	
96. <i>Other</i>	
97. <i>Other</i>	
98. <i>Other</i>	
99. <i>Other</i>	
100. <i>Other</i>	

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4703159"
/clone_lib="NIH_MGC_77"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: lung; Vector: pDNR-LIB (Clontech); Site 1: SfiI (ggccgcctgcgc); Site 2: SfiI (ggccattatgcc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGCCATTATGCCC-3' and 3' adaptor sequence: 5'-ATTCTAGCGGCGGCGGCCGACATC-dT(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.9 kb (range 0.5-4.0 kb). 12/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH_MGC Library."
113 a 202 c 163 q 120 t

```

BASE COUNT
ORIGIN

113 a 202 c 163 g 120 t

Note: this is a NIMMGC Library.

29.9% Score 312.8; DB 11; Length 598;
Similarity 99.4%;
Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Query Match	29.9%	Score 312.8;	DB 11;	Length 598;
Best Local Similarity	99.4%;	Pred. No. 8.3e-61;		
Matches 314;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;
QY 732	GCACCTGGCAGCTGCTTGAGCCAACTCCATAGCGGCTTTTCTCTACACAGGGGCTCCT	791		
Db				
185	GAAGCTGGCAGGCTGCTTGAGCCAACTCCATAGCGGCTTTTCTCTACACAGGGGCTCCT	244		
QY 792	GCAGGCGCTGGAAGGGGATCTCCCCGAGTTGGGTGCCACCTTGGACACACTGCAGCTGGA	851		
Db				
245	GCAGGCGCTGGAAGGGGATCTCCCCGAGTTGGGTGCCACCTTGGACACACTGCAGCTGGA	304		
QY 852	CGTCCGCGACTTTGCCACCACTATCTGGCAGCAGATGGGAAGAACTGGGAATGCCCTGC	911		
Db				
305	CGTCCGCGACTTTGCCACCACTATCTGGCAGCAGATGGGAAGAACTGGGAATGCCCTGC	364		
QY 912	CTGTGACGCCACCAAGGGTGCCATCGCGGCTTCGTGCTTCACGCGCGGCGAG	971		
Db				
365	CTGTGACGCCACCAAGGGTGCCATCGCGGCTTCGTGCTTCACGCGCGGCGAG	424		
QY 972	AGGGGTCTAGTTGGCTTCCCATCTGCAGAGCTTCCTGAGGTTGCTCTACCGGTTCTACG	1031		
Db				
425	AGGGGTCTAGTTGGCTTCCCATCTGCAGAGCTTCCTGAGGTTGCTCTACCGGTTCTACG	484		
QY 1032	CCACCTTGCCCAAGCCC	1047		
Db				
485	CCACCTTGCCCAAGCCC	500		

RESULT 3
BT328788

RESULT	3
BI328788	
LOCUS	BI328788 900 bp mRNA EST 30-JUL-2001
DEFINITION	602984776F1 NCI_CGAP_L19 Mus musculus cDNA clone IMAGE:5137578 5', mRNA sequence.
ACCESSION	BI328788
VERSION	BI328788.1 GI:15013445
KEYWORDS	EST.
SOURCE	house mouse.
ORGANISM	Mus musculus
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 900)
AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: LLAM11338 row: 9 column: 19
High quality sequence stop: 808.

FEATURES
source
1. .900
Location/Qualifiers
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:5137578"
/clone_lib="NCI_CGAP_Li9"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: liver; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.9 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 194 a 290 c 223 g 192 t 1 others
ORIGIN

Query Match 28.8%; Score 301.6; DB 11; Length 900;
Best Local Similarity 85.1%; Pred. No. 3.1e-58;
Matches 337; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 1 ATGAGCTGACTGAATGCTCCTCGTGTATGCTTCTCTAACTGCAAGCTAACGCTG 60
|||||
DB 191 ATGAGCTGACTGATTTGCTCTCGGGGCCATGCTTTCAGTGGCAAGACTAATCTG 250
QY 61 TCAGCGCGGCTCTCTGCTGTGAGCTCCGAGTCTCAGTAACTGCTGTGACTCC 120
|||||
DB 251 TCAGCGCGGCTCTCTGCTGTGAGCTCCGAGTCTCAGTAACTGCTGTGACTCC 310
QY 121 CATGCTCTTCACAGCAGTACGAGTCCGAGGTTTCCACCTTTGCTTACACCTGTC 180
|||||
DB 311 CACCTCTTCACAGCAGTACGAGTCCGAGGTTTCCACCTTTGCTTACACCTGTC 370
QY 181 CTGCTGCTCTGTGGACTTTAGCTTGGGAGATGGAAACCCAGATGGAGGACCAAG 240
|||||
DB 371 CTGCTGCTCTGTGGACTTTAGCTTGGGAGATGGAAACCCAGATGGAGGACCAAG 430
QY 241 GCACAGACATTCGGAGCAGTACACCTTCTGCTGAGGAGTGTGGCAGCAGCGGA 300
|||||
DB 431 GCACAGACATTCGGAGCAGTACACCTTCTGCTGAGGAGTGTGGCAGCAGCGGA 490
QY 301 CAATGGGACCCACTTGCCTCTCATCTCTCTGCGGAGCTTTCTGACAGGTCCTC 360
|||||
DB 491 CAGTTGGAACCCCTCTGCTCTCTCATCTCTCTGCGGAGCTTTCTGAGCAGGTTTCGCTC 550
QY 361 CTCTTGGGCGCCTGACAGCCTCTTGGAAACCCAG 396
|||||
DB 551 CTCTTGGGCGCCTGACAGCCTCTTGGAAACCCAG 586

RESULT 4
LOCUS BE485194 507 bp mRNA EST 28-AUG-2000
DEFINITION 172130 BARC 580V Bos taurus CDNA 5', mRNA sequence.
ACCESSION BE485194
VERSION BE485194.1 GI:9604727
KEYWORDS EST.
SOURCE COW.

ORGANISM

Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 507)
REFERENCE AUTHORS Sonstegard,T.S., Capuco,A.V., Van Tassel,C.P., Ashwell,M.S. and Wells,K.D.

TITLE Mapping of Expressed Sequence Tags from a normalized bovine mammary gland cDNA library
JOURNAL Unpublished (2000)
COMMENT Contact: Sonstegard TS
USDA, ARS, Beltsville Agricultural Research Center
Bldg. 200 Rm 2A, Beltsville, MD 20705, USA
Tel: 301 504 8416
Fax: 301 504 8414
Email: tads@psi.barc.usda.gov

Single pass sequencing. Bases called and alt_trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -minmatch 12 options.

PCR Primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCAGTCACGACG
Plate: 134 row: J column: 17
Seq primer: ATTAGGTGACACTATAG.
Location/Qualifiers

FEATURES

source
1. .507
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone_lib="BARC 580V"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: pCMV SPORT6; Site 1: XbaI; Site 2: XhoI;
Library made from pooled mRNA isolated from mammary
tissues at eight physiological, developmental, and disease
states."
BASE COUNT 97 a 170 c 158 g 82 t
ORIGIN

Query Match 28.7%; Score 300.6; DB 10; Length 507;
Best Local Similarity 84.1%; Pred. No. 4.6e-58;
Matches 339; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 630 GTGTGCCACCTACAAGCTGTGCCACCCGAGGAGCTGCTGCTCGACACTCTCTGGG 689
|||||
DB 104 GTGTGCCGCCCAAGCTGTGCCACCCGAGGAGCTGCTGCTCAGGCACTCTCTGGG 163
QY 690 CATCCCTTGGGCTCCCTGAGCAGCTGCCCCAGGCGCTCGAGCTGCGAGGCTGCTT 749
|||||
DB 164 CATCCCTTGGGCTCCCTGAGCAGCTGCTCCAGCAGTCCCTGCAGCTGACGAGCTGCT 223
QY 750 GAGCCAACTCCATAGCGGCTTTTCTCTACAGGGGCTCTTCAGGCGCTTGGAGGGAT 809
|||||
DB 224 GAACCAACTACAGCGGCGCTCTTCTCTACAGGGGCTCTTCAGGCGCTTGGAGGGAT 283
QY 810 CTCCCGGAGTGTGGTCCCACTTGGACACTTGGACACTGACGCTGCGGAGCTTGGCCAC 869
|||||
DB 284 CTCCCGGAGTGTGGTCCCACTTGGACACTTGGACACTGACGCTGACGCTTGGCCAC 343
QY 870 CACCATCTGGCAGCAGATGGAAGACTTGGGAATGGCCCTTGCCTTGCAGCCACCCAGG 929
|||||
DB 344 GAACATCTGGCTGCAGATGGAGACCTGGGGCGGCGGCGCTGCTGAGGCGGCGGCGG 403
QY 930 TGCATGCGCGGCTTGGCTCTGCTTTCCAGGCGGCGGCGGAGGAGGCTCTAGTTGCTC 989
|||||
DB 404 CGCCATGCCGACCTTCACTTCAGCCTTCCAAACGACAGGAGGAGGCTCTGTTGCTTC 463
QY 990 CCATCTGCAGAGCTTCTCTGAGGTGCTGTACCGCTTCTACGC 1032
|||||
DB 464 CCAGCTGCATCGTTTCTTGAGCTGGCATACCGTGGCTCGGC 506

RESULT 5

```

AI1314551      633 bp      mRNA      EST      17-DEC-1998
LOCUS          u148c10.y1 Sugano mouse liver mlia Mus musculus cDNA clone
DEFINITION     IMAGE:1923186 5' similar to gb:L34169 Mus musculus thrombopoietin
               mRNA, complete cds (MOUSE);, mRNA sequence.
ACCESSION      AI1314551
VERSION        AI1314551.1 GI:4029818
KEYWORDS       EST.
SOURCE         house mouse.
ORGANISM       Mus musculus.
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE      1 (bases 1 to 633)
AUTHORS        Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
               Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
               Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
               Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
               Waterston,R.
TITLE          The WashU-HMI Mouse EST Project
JOURNAL        Unpublished (1996)
COMMENT        Contact: Marra M/Mouse EST Project
               WashU-HMI Mouse EST Project
               Washington University School of Medicine
               4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
               Tel: 314 286 1800
               Fax: 314 286 1810
               Email: mouseest@watson.wustl.edu
               This clone is available royalty-free through LLNL; contact the
               IMAGE Consortium (info@image.llnl.gov) for further information.
               MGI:979478
FEATURES       Seq primer: custom primer used
               High quality sequence stop: 505.
               Location/Qualifiers
               1..633
               /organism="Mus musculus"
               /strain="C57BL"
               /db_xref="taxon:10090"
               /clone="IMAGE:1923186"
               /clone_lib="Sugano mouse liver mlia"
               /sex="female"
               /dev_stage="adult"
               /lab_host="DH10B"
               /note="Organ: liver; Vector: pME18S-FL3; Site_1: DraIII
               (CACTGTGTG); Site_2: DraIII (CACCATGTG); 1st strand cDNA
               was primed with an oligo(dT) primer
               [ATGTGGCTTTTCTTTTCTTTT]; double-stranded cDNA was
               ligated to a DraIII adaptor [TGTGGCTACTGG], digested
               and cloned into distinct DraIII sites of the pME18S-FL3
               vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should
               be used to isolate the cDNA insert. Size selection was
               performed to exclude fragments <1.5kb. Library
               constructed by Dr. Sumio Sugano (University of Tokyo
               Institute of Medical Science). Custom primers for
               sequencing: 5' end primer CTCTGTGCTCTAAAGCTGG and 3' end
               primer GCACCTGCAGCTCGAGACA."
BASE COUNT     125 a 210 c 166 g 131 t 1 others
ORIGIN
Query Match      28.7%; Score 300.6; DB 10; Length 633;
Best Local Similarity 84.8%; Pred. No. 4.9e-58;
Matches 336; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 1 ATGAGCTGACTGAATGTCTCTGTGGTCAATGCTTCTTAACTGCAAGGTAACGCTG 60
|||||
Db 238 ATGAGCTGACTGATTTGCTCTCGGGCCATGCTCTGTGAGTGGCAAGACTACTCTG 297
|||||

QY 61 TCACGCCGCTCTCTCTGTGTGACCTCGAGTCCCTCAGTAACTGCTGTGACTGCC 120
|||||
Db 298 TCACGCCGCTGCTCTCTGCTGTGACCCAGACTCTAAATAAACTGCTGTGACTGCC 357
|||||

QY 121 CATGTCTTCACAGCAGACTGAGCAGTGCAGAGGCTTACCCCTTTCCTACACCTGTC 180
|||||

```

```

Db 358 CACCTCTTCACAGCCGACTGAGTCAGTGTCCCGAGCTCGACCCCTTTGTCTATCCCTGTT 417
QY 181 CTGCTCCCTGCTGTGAGCACTTTAGCTTTGGGAGAATGAAAACCCAGATGGAGGACCAAG 240
|||||
Db 418 CTGCTCCCTGCTGTGAGCACTTTAGCTTTGGGAGAATGAAAACCCAGAGGACGAGCAAG 477
|||||
QY 241 GCACAGGACATTTCTGGGAGCAGTACCTTCTCTCTCTGGAGGAGTGTATGGCAGCAGGGGA 300
|||||
Db 478 GCACAGGACATTTCTAGGGGAGTGTCCCTTCTACTTGAGGGAGTGTATGGCAGCAGGAGGA 537
|||||
QY 301 CAACTGGGAGCCACTTGCCTCTCATCTCCCTCGGGGAGAGCTTTTCTGGACAGGTCCGCTC 360
|||||
Db 538 CAGTTTGAACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 597
|||||
QY 361 CTCTCTGGGGCCCTGCGAGAGCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 396
|||||
Db 598 CTCTCTGGGGCCCTGCGAGAGCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 633
|||||

RESULT        6
LOCUS          AA538257      609 bp      mRNA      EST      29-JUL-1997
DEFINITION     VJ03a07.r1 Barstead mouse pooled organs MPLRB4 Mus musculus cDNA
               clone IMAGE:920628 5' similar to gb:L34169 Mus musculus
               thrombopoietin mRNA, complete cds (MOUSE);, mRNA sequence.
ACCESSION      AA538257
VERSION        AA538257.1 GI:2284250
KEYWORDS       EST.
SOURCE         house mouse.
ORGANISM       Mus musculus.
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE      1 (bases 1 to 609)
AUTHORS        Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
               Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
               Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
               Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
               Waterston,R.
TITLE          The WashU-HMI Mouse EST Project
JOURNAL        Unpublished (1996)
COMMENT        Contact: Marra M/Mouse EST Project
               WashU-HMI Mouse EST Project
               Washington University School of Medicine
               4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
               Tel: 314 286 1800
               Fax: 314 286 1810
               Email: mouseest@watson.wustl.edu
               This clone is available royalty-free through LLNL; contact the
               IMAGE Consortium (info@image.llnl.gov) for further information.
               MGI:532844
               Seq primer: -28ml3 rev2 ET from Amersham
               High quality sequence stop: 492.
               Location/Qualifiers
               1..609
               /organism="Mus musculus"
               /strain="FVB/N"
               /db_xref="taxon:10090"
               /clone="IMAGE:920628"
               /clone_lib="Barstead mouse pooled organs MPLRB4"
               /sex="mixed"
               /tissue_type="pooled organs"
               /dev_stage="7 day"
               /lab_host="DH10B"
               /note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
               a modified polylinker; Site_1: EcoRI; Site_2: NotI; 1st
               strand cDNA was primed with a Not I - oligo(dT) primer [5'
               TGTACGAATCTGAGTGGAGCGGCCCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT
               3']; double-stranded cDNA was ligated to Eco RI adaptors
               [GTTGGATTGGTACC], digested with Not I and cloned into
               the Not I and Eco RI sites of the modified pT7T3 vector.
               Library constructed by Bob Barstead."
BASE COUNT     126 a 204 c 157 g 122 t
ORIGIN

```

1:

RESULT	7
BF789308	
LOCUS	878 bp mRNA EST 12-JAN-2001 602104729F1 NCI_CGAP_Kid14 Mus musculus cDNA clone IMAGE:4222927
DEFINITION	5', mRNA sequence.
ACCESSION	BF789308
VERSION	BF789308.1 GI:12094344
KEYWORDS	EST.
SOURCE	house mouse.
ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Rodentia; Muridae; Murinae; Mus;
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/.
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov Tissue Procurement: Jeffrey E. Green, M.D. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM9810 row: 1 column: 08 High quality sequence stop: 628. Location/Qualifiers 1. -878 /organism="Mus musculus" /strain="FVB/N" /db_xref="taxon:10090" /clone="IMAGE:4222927" /clone.lib="NCI CGAP Kid14"
FEATURES	
SOURCE	

1:

RESULT	8	
BE754587		
LOCUS	464 bp	EST
DEFINITION	208137 MARC 2BOV Bos taurus cDNA 5', mRNA sequence.	25-APR-2001
ACCESSION	BE754587	
VERSION	BE754587.1	GI:10168579
KEYWORDS	EST.	
SOURCE	cow.	
ORGANISM	Bos taurus	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidae;	
	Bovidae; Bovinae; Bos.	
	1 (bases 1 to 464)	
REFERENCE	Smith,E.P.L., Grose,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,	
AUTHORS	Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett	
	G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C.G.,	
	Perlea,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and	
	Keefe,J.W.	
TITLE	Sequence evaluation of four pooled-tissue normalized bovine CDNA	
JOURNAL	libraries and construction of a gene index for cattle	
MEDLINE	Genome Res. 11 (4), 626-630 (2001)	
COMMENT	21180013	
	Contact: Smith TPL	
	USDA, ARS, US Meat Animal Research Center	
	PO Box 166, Clay Center, NE 68933-0166, USA	
	Tel: 402 762 4366	
	Fax: 402 762 4390	
	Email: smith@email.marc.usda.gov	
	Single pass sequencing. Bases called and alt_trimmed with phred	
	v0.980904.e. Vector identified by cross_match with the -minscore 18	
	and -mismatch 12 options.	

PCR Primers

FORWARD: AGGAACAGCTATGACCAT
 BACKWARD: GTTTCCTCAGCAGCAG
 Plate: 54 row: J column: 23
 Seq primer: ATTAGTGACACTATAG.

FEATURES

source

1..464
 /organism="Bos taurus"
 /db_xref="taxon:9913"
 /clone_lib="MARC 2Bov"
 /tissue_type="pooled"
 /lab_host="DH10B"
 /note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
 Library made from pooled tissue from testis, thymus,
 semitendinosus muscle, longissimus muscle, pancreas,
 adrenal, and endometrium."
 97 a 152 c 128 g 87 t

BASE COUNT

ORIGIN

Query Match 22.0%; Score 230; DB 10; Length 464;
 Best Local Similarity 85.5%; Pred. No. 4.3e-42;
 Matches 284; Conservative 0; Mismatches 35; Indels 13; Gaps 2;
 QY 203 GCTTGGGAGATGGAACCCAGATGGAGGAGACCAAGGACAGACATTCCTGGAGCAG 262
 Db 1 GCTTGGGAGATGGAACCCAGATGGAGGAGACCAAGGAGACATTCCTGGAGCAG 262
 QY 263 TGACCCCTTCCTGGAGGAGTGGGAGCAGCGGGGACAACTGGACCCCTCT 322
 Db 61 CGACCCCTTCCTGGAGGAGTGGGAGCAGCGGGGACAACTGGACCCCTCT 322
 QY 323 CATCCCTCTGGGAGCTTTCGGACAGTCCGTCCTCTCTTGGGGCCCTGCAGAGC 382
 Db 121 CATCCCTCTGGGAGCTTTCGGAGGTCGCGCTCTCTTGGGGCCCTGCAGAGC 382
 QY 383 TCCTTGAACCCAGCTTCTCCACAGGCGAGGACACAGCTCACAAGGATCCCAATGCCA 442
 Db 181 TCCTGAACCC-----CAGGCGAGGACACAGCTCACAAGGATCCCAATGCCA 442
 QY 443 TCTTCTGAGCTTCCAAACACCTGCTCCGAGGAAAGTGGCTTCTCTGATGCTGT-AGGA 501
 Db 229 TCTTCTGAGCTTCCAAACAGCTGCTCCGAGGAAAGTGGCTTCTCTGATGCTGT-AGGA 501
 QY 502 GGGTCCACCTCTCGTACGCGGGGCCCAAC 533
 Db 289 GGGCCACCTCTGTGCAAGCGGGGCCCAAC 320

RESULT 9

BF848766

LOCUS

BF848766 327 bp mRNA

QV0-EN0102-081100-458-g06 EN0102 Homo sapiens EST 16-JAN-2001

ACCESSION BF848766 QV0-EN0102-081100-458-g06 EN0102 Homo sapiens cDNA, mRNA sequence.

VERSION BF848766.1 GI:12235903

KEYWORDS EST.

SOURCE human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 327)

AUTHORS Dias Neto,B., Garcia Correa,R., Verjovsky-Almeida,S., Briones,M.R.,

Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,

Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,

Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare

,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and

Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

PROC. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

JOURNAL 20202663

CONTACT: Simpson A.J.G.

COMMENT Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl-QV0&t2-QV0-EN0102-
 081100-458-g06&t3-2000-11-08&t4-1)

Seq primer: puc 18 forward

High quality sequence stop: 327.

FEATURES

source

1..327
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="EN0102"
 /dev_stage="Adult"

/note="Organ: lung, normal; Vector: puc18; Site_1: SmaI;
 Site_2: SmaI; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the pUC 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."
 61 a 87 c 103 g 76 t

BASE COUNT

ORIGIN

Query Match 13.0%; Score 135.8; DB 11; Length 327;
 Best Local Similarity 89.6%; Pred. No. 8.4e-21;
 Matches 146; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
 QY 726 GGCCCTGCAGCTGGCAGCGCTGCTTGAGCCAACTCCATAGCGCCCTTTCTCTACCAAGG 785
 Db 139 GGCCCATACACAGCGCGCGCTTTTGACCACTCCATAGCGCCCTTTCTCTACCAAGG 785
 QY 786 GCTCCTGCAGCCCTGGAAGGATCTCCCGAGTTGGTCCCACTTGGACACACTGCA 845
 Db 199 GCTCCTGCAGCCCTGGAAGGATCTCCCGAGTTGGTCCCACTTGGACACACTGCA 258
 QY 846 GCTGAGCTGCGGACTTTGCCACCACTCTGCGACGAGTG 888
 Db 259 GCTGAGCTGCGGACTTTGCCACCACTCTGCGACGAGTG 301

RESULT 10

AZ025993

LOCUS

AZ025993 449 bp DNA

RPCI-23-343116.TJ RPCI-23 Mus musculus genomic clone RPCI-23-343116

DEFINITION , DNA sequence.

ACCESSION AZ025993

VERSION AZ025993.1 GI:7101377

KEYWORDS GSS.

SOURCE house mouse.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

REFERENCE 1 (bases 1 to 449)

AUTHORS Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S., Akinret

,B., Levins,M., Mcgann,S., Isegaye,G., Geer,K., Krol,M., de Jong,P.

and Fraser,C.M.

Mouse BAC End Sequences from Library RPCI-23

Unpublished (1999)

JOURNAL Other_GSSs: RPCI-23-343116.TV

COMMENT Contact: Shaying Zhao

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: szhao@tigr.org

Clones are derived from the mouse BAC library RPCI-23. For BAC

library availability, please contact Pieter de Jong (pieter@jeong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (<http://bacpac.med.buffalo.edu/orderingframe.htm>) or from Resea ch Genetics (info@resgen.com). BAC end page: http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
 Plate: 343 row: I column: 16
 Seq primer: SP6
 Class: BAC ends.

FEATURES

source
 1..449
 Location/Qualifiers
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="RPC1-23-343I16"
 /clone_lib="RPC1-23"
 /sex="Female"
 /lab_host="DH10B"
 /note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1: ECORI; Site_2: ECORI; Female C57BL/6J mouse kidney and/or brain genomic DNA was isolated and partially digested with a combination of EcoRI and EcoRI Methyase. Size selected DNA was cloned into the pBACe3.6 vector at the ECORI sites. The ligation products were transformed into DH10B electrocompetent cells (BRL Life Technologies)."

BASE COUNT 90 a 121 c 108 g 129 t 1 others

ORIGIN

Query Match 12.9%; Score 135.2; DB 13; Length 449;
 Best Local Similarity 84.4%; Pred. No. 1.2e-20;
 Matches 152; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY 228 GGAGGAGCAGGAGCAGGAGCAGTCTGGGAGCAGTACCCCTCTGCTGGAGGAGTGCAT 287
 DB 126 GGAACAGCAGGAGCAGGAGCAGTCTGGGAGCAGTCTGCTGGAGGAGTGCAT 185

QY 288 GGCAGCAGGAGCAGGAGCAGGAGCAGTCTGCTCTCATCCCTCTGGGAGGAGTCTGG 347
 DB 186 GGCAGCAGGAGCAGGAGTGGAAACCTCTGCTCTCATCCCTCTGGGAGGAGTCTGG 245

QY 348 ACAGGTCGGTCTCTCTCTGGGAGGAGCAGGAGCAGTCTGGAAACGAGTCTCCACACA 407
 DB 246 GCAGGTTGGCTCTCTCTGGGAGGAGCAGGAGCAGTCTGGAAACGAGTCTCCACACA 305

RESULT 11

AZ352219/c
 LOCUS 419 bp DNA GSS 29-SEP-2000
 DEFINITION 1M0090P05R Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0090P05 R, DNA sequence.
 ACCESSION AZ352219
 VERSION AZ352219.1 GI:10431456
 KEYWORDS GSS.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 419)
 AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
 TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
 JOURNAL Unpublished (2000)
 COMMENT Contact: Robert B. Weiss
 University of Utah Genome Center
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00
 Plate: 0090 row: P column: 05
 Seq primer: CACACAGGAACACCTATGACC
 Class: plasmid ends
 High quality sequence stop: 419.
 Location/Qualifiers

FEATURES

source
 1..419
 Location/Qualifiers
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGC1M0090P05"
 /clone_lib="Mouse 10kb plasmid UUGC1M library"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
 /note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PWD42 (gil4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 100 a 98 c 124 g 97 t

ORIGIN

Query Match 10.4%; Score 109.4; DB 13; Length 419;
 Best Local Similarity 76.6%; Pred. No. 8.4e-15;
 Matches 134; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 360 CCTCTTGGGGCCCTCGAGAGCCTCTTGGAAACCCAGCTTCTCTCACAGGCGAGGACCAC 419
 DB 292 CCTACTCAGTGGCTCAGGCGCCATCCTCTGCCCTCAGCTTCTCTACAGGCGAGGACCAC 233

QY 420 AGCTCACAAGATCCCAATGCCATCTTCTCAGCTTCCACACCTGCTCCGAGGAAAGGT 479
 DB 232 AGCTCACAAGAGCCCAATGCCCTCTTCTTTCAGCTTCCAACTGCTTCCGGGAAAGGT 173

QY 480 GCGTTTCTGATGCTTCTAGGAGGTCACCCCTCTGCTAGCGGGCGGCGGCAACA 534
 DB 172 GCGCTTCTGCTTCTGGTAGAAGTCCACCCCTCTGCTGTCAGAGGACCCCTGCCA 118

RESULT 12

AZ792988
 LOCUS 708 bp DNA GSS 16-FEB-2001
 DEFINITION 2M0046I05F Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC2M0046I05 F, DNA sequence.
 ACCESSION AZ792988
 VERSION AZ792988.1 GI:12937676
 KEYWORDS GSS.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 708)
 AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
 TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts

JOURNAL
COMMENT

Unpublished (2000)
Contact: Robert B. Weiss
University of Utah
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: rdunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0046 row: 1 column: 05
Seq primer: CGTGTAAACGACGCCAGT
Class: plasmid ends
High quality sequence stop: 708.
Location/Qualifiers
1. .708
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0046105"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/notes="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pWD42 (g1/4732114/gb/AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

FEATURES

source

1. .708
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0046105"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/notes="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pWD42 (g1/4732114/gb/AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT
ORIGIN

156 a 213 c 147 g 192 t

Query Match

Best Local Similarity 8.3%; Score 86.4; DB 13; Length 708;
Matches 111; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

Qy 16

TTGCTCTCGTGGTGCATGCTTCTCCCTAACGCAAGGCTAACGCTGCCAGCCGGCTCCT 75

Db 1

TTGCTCTCGCGCCATGCTTCTTGCAAGTGGCAAGACTAACTCTGTCCAGCCCGTAGCT 60

Qy 76

CCTGCTTGTGACCTCGGAGTCTCAGTAAACTGCTTCGTGACTCCCATGCTTCCACAGC 135

Db 61

CCTGCTTGTGACCTCGGAGTCTCAGTAAACTGCTTCGTGACTCCCATGCTTCCACAGC 120

Qy 136

AGACTGAGCCAGTGCACAGAGTTTACCCCTTT 167

Db 121

CGACTGCTGAGCAACCCCAAGCCCTCGCTTT 152

RESULT 13

AQ888740

LOCUS

AQ888740 752 bp DNA GSS 10-NOV-1999

DEFINITION

HS_3106_B1_B04_MR CIT Approved Human Genomic Sperm Library D Homo

ACCESSION

AQ888740 sapiens genomic clone Plate=3106 Col=7 Row=D, DNA sequence.

VERSION

AQ888740.1 GI:6344930

KEYWORDS

GSS.

SOURCE

human.

ORGANISM

Homo sapiens

REFERENCE

AUTHORS
Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T., and
Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D., and
Hood, L.

TITLE

Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome

JOURNAL

Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)

MEDLINE

99380589

COMMENT

Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones may be purchased from Research Genetics (info@resgen.com).
BAC end Web Server: http://www.htsc.washington.edu
Plate: 3106 row: D column: 7
Seq primer: M13 Reverse
Class: BAC ends
High quality sequence stop: 752.

FEATURES

source

1. .752
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Plate=3106 Col=7 Row=D"
/clone_lib="CIT Approved Human Genomic Sperm Library D"
/sex="male"
/note="Organ: sperm; Vector: pBelobAC11; BAC Clones in
E-Coli DH10B"
BASE COUNT 181 a 219 c 177 g 175 t
ORIGIN

Query Match

Best Local Similarity 7.4%; Score 77.8; DB 13; Length 752;
Matches 123; Conservative 0; Mismatches 57; Indels 1; Gaps 1;

Qy 13

GAATTCCTCTCGTGGTGCATGCTTCTCCCTAACGCAAGGCTAACGCTGCCAGCCCGGCT 72

Db 509

GAATTCCTCTCGTGGTGCATGCTTCTCCCTAACGCAAGGCTAACGCTGCCAGCCCGGCT 568

Qy 73

CCTCCTCTCTGACCTCGGAGTCTCAGTAAACTGCTCGTGCATCCCATGCTTCTCAC 132

Db 569

CCTACTGTTGTGAACCTCCATCTCAG-GAAGTGTTCGGGACTCCCATGGCTCACA 627

Qy 133

AGCAGACTGAGCCAGTGCAGAGGTTTCCACCTTTGCCCTACACCTGTCCTGCTGCT 192

Db 628

GCAAACTGGTGAGATCTCCAAGTATCCCTTTATCCGCTTAAGTGTAAAGACCATACT 687

Qy 193

G 193

Db 688

G 688

RESULT 14

BE377769

LOCUS

BE377769 589 bp mRNA EST 21-JUL-2000

DEFINITION

60122920F1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:3593885 5',
mRNA sequence.

ACCESSION

BE377769

VERSION

BE377769.1 GI:9323134

KEYWORDS

EST.

SOURCE

house mouse.

ORGANISM

Mus musculus

REFERENCE

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

AUTHORS

1 (bases 1 to 589)

TITLE

NIH-MGC http://mgs.nci.nih.gov/

JOURNAL

National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.
Email: cchapb-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM8767 row: g column: 06
High quality sequence stop: 587.

FEATURES

source
1. 589
Location/Qualifiers
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone_image="393885"
/clone_lib="NCI-CGAP_Maml"
/tissue_type="tumor, biopsy sample"
/dev_stage="3 months, virgin"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site: Sali;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"

BASE COUNT

150 a 152 c 134 g 153 t

Query Match

Best Local Similarity 6.3%; Score 65.6; DB 10; Length 589;
Matches 80; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

Qy 1 ATGAGCTGACTGATTCCTCGTGTGATGCTCTCTCACTGCAAGCTCAAGCTG 60
|||||
Db 484 ATGAGCTGACTGATTCCTCGTGTGATGCTCTCTCACTGCAAGCTCAAGCTG 543
|||||
Qy 61 TCCAGCCCGGCTCTCTCGTGTGACCTCGGAGCTCTCTCACTGCAAGCTG 104
|||||
Db 544 TCCAGCCCGGCTCTCTCGTGTGACCTCGGAGCTCTCTCACTGCAAGCTG 587
|||||

RESULT 15

CNS0091P 925 bp DNA GSS 03-JUN-1999
LOCUS Drosophila melanogaster genome survey sequence TET3 end of BAC #
DEFINITION BACR19D16 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.

ACCESSION

AL053013.1 GI:4934461

VERSION

GSS.

KEYWORDS

fruit fly.

ORGANISM

Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;

Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

Genoscope.

Direct Submission

Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :

BP 191 91006 EVRY cedex - FRANCE (E-mail : secref@genoscope.cns.fr)

- Web : www.genoscope.cns.fr

Determination of this BAC-end sequence was carried out as part of a

collaboration with the Berkeley Drosophila Genome Project (BDGP).

The BDGP is constructing a physical map of the Drosophila

melanogaster genome using these BACs. For further information

please see <http://www.fruitfly.org> The BDGP Drosophila

melanogaster BAC library was prepared by Kazutoyo Osogawa and

Aaron Mammoser in Pieter de Jong's laboratory in the Department of

Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,

NY. The library is named RPCI-98 and was constructed by partial

EcoRI digestion of Drosophila DNA provided by the BDGP from the

isogenic strain y2; cn bw sp, the same strain used for the BDGP's

PI and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

source
1. 925
Location/Qualifiers
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_lib="RPCI-98"
/clone="BACR19D16"
/note="end : TET3"

BASE COUNT

120 a 61 c 172 t 511 others

ORIGIN

Query Match 5.2%; Score 54.8; DB 13; Length 925;
Best Local Similarity 15.5%; Pred. No. 0.023;
Matches 66; Conservative 174; Mismatches 182; Indels 3; Gaps 1;
Qy 555 ACAGAGCTTCCTTTTAAAGCTTAGAGCAAGTGAAGAGATCCAGGCGGATGGCGCAGC 614
|||||
Db 492 ANANNNTTATTATTANNNNNANANANANNNNNAGCSMSMCKKCGSTTBGSTTT 551
|||||
Qy 615 GCTCCAGGAGAGCTGTGTCACCTACAAGCTGTGCCACCCGAGGAGCTGGTGCTGCT 674
|||||
Db 552 TTTSSGGYGRGCSGSGSCSCSCSCSCSCSCSCSCSCSCSCSCSCSCSCSCSCSCSCS 611
|||||
Qy 675 CGGACACTCTCTGGGCATCCCTGGCTCCCTGAGCAGCTGCCCCAGCCAGCCCTGCA 734
|||||
Db 612 STSBSCSCSSKSVCGTSCSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 671
|||||
Qy 735 GCTGCGAGCTGCTGAGCCCACTCATAGCGGCTTTTCTCTACCAAGGCTCTCTGCA 794
|||||
Db 672 SKSTASGSGSWAGGSGSTGTSSTSSSSSSSTSTSSSSSSSSSSSSSSSSSSSSSS 731
|||||
Qy 795 GGCCCTGGAAGGATCTCCCGGATGGTCCACACTTGGACACACTGCA---GCTGGA 851
|||||
Db 732 SSSSTSSBBSCTSTSSSSSSSYSSSTCCCTCCGCSYSSSTSSSSSTSSSSSSSS 791
|||||
Qy 852 CGTCCCGACTTTCACACCACTATGCGACAGATGGAAGAACTGGGAATGGCCCTGC 911
|||||
Db 792 VGTSSSDSTSTCCSCCCYMCCTCCSTVBMBCYTSTSCGSSSSSSGKGVTKCGCGGCGSS 851
|||||
Qy 912 CCGCAGCCACCCAGGCTGCCATCGCGCTTCCTCTCTTTCAGCGCCGCGGAGG 971
|||||
Db 852 STNGMBGTSSACSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 911
|||||
Qy 972 AGGGG 976
|||||
Db 912 VSSGS 916
|||||

Search completed: April 5, 2002, 09:47:00
Job time: 3324 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 5, 2002, 10:29:29 ; Search time 2520.15 Seconds
(without alignments)
7089.437 Million cell updates/sec

Title: US-09-680-514-6

Perfect score: 1083

Sequence: 1 ATGGAGCTGACTGAATTGCT.....TACGCCACCTTGCCAGCCC 1083

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*

- 1: gb_ba:*
- 2: gb_htg:*
- 3: gb_in:*
- 4: gb_om:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_sts:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vi:*
- 15: em_ba:*
- 16: em_fun:*
- 17: em_hum:*
- 18: em_in:*
- 19: em_om:*
- 20: em_or:*
- 21: em_ov:*
- 22: em_pat:*
- 23: em_ph:*
- 24: em_pl:*
- 25: em_ro:*
- 26: em_sts:*
- 27: em_sy:*
- 28: em_un:*
- 29: em_vi:*
- 30: em_htgo_hum:*
- 31: em_htgo_inv:*
- 32: em_htgo_rod:*
- 33: em_htg_hum:*
- 34: em_htg_inv:*
- 35: em_htg_rod:*
- 36: em_htg_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Query	Description
1	523.2	48.3	555	22	E11965		E11965 Human cDNA
2	523.2	48.3	1059	6	AR103959		AR103959 Sequence
3	523.2	48.3	1062	6	AR087133		AR087133 Sequence
4	523.2	48.3	1062	6	AR087571		AR087571 Sequence
5	523.2	48.3	1062	6	I49760		I49760 Sequence 3
6	523.2	48.3	1062	9	HUMTHROMB		L36052 Human throm
7	523.2	48.3	1086	6	E16669		E16669 cDNA encodi
8	523.2	48.3	1267	6	E12182		E12182 Human cDNA
9	523.2	48.3	1267	6	E12214		E12214 Human cDNA
10	523.2	48.3	1342	6	AR008878		AR008878 Sequence
11	523.2	48.3	1342	6	AR023468		AR023468 Sequence
12	523.2	48.3	1342	6	AR087573		AR087573 Sequence
13	523.2	48.3	1342	6	I85129		I85129 Sequence 1
14	523.2	48.3	1691	9	D32047		D32047 Homo sapien
15	523.2	48.3	1721	6	E16668		E16668 cDNA encodi
16	523.2	48.3	1721	22	E11961		E11961 Human cDNA
17	523.2	48.3	1795	6	I33525		I33525 Sequence 1
18	523.2	48.3	1795	9	HUMMLCMLP		L33410 Human c-mpl
19	522	48.2	525	6	E02574		E02574 DNA encodin
20	521.6	48.2	605	6	AR008879		AR008879 Sequence
21	521.6	48.2	605	6	AR087574		AR087574 Sequence
22	521.6	48.2	605	6	I85130		I85130 Sequence 3
23	520	48.0	1341	9	HSU11025		U11025 Human megak
24	518.4	47.9	1062	9	HSU59493		U59493 Human throm
25	518.4	47.9	1062	9	HSU59494		U59494 Human throm
26	515.6	47.6	525	6	E02575		E02575 DNA encodin
27	514	47.5	525	6	E02576		E02576 DNA encodin
28	489.2	45.2	1050	22	E11559		E11559 Human cDNA
29	486.8	44.9	525	6	AR024358		AR024358 Sequence
30	486.8	44.9	525	6	AR091731		AR091731 Sequence
31	486.8	44.9	525	6	E01731		E01731 cDNA encodi
32	486.8	44.9	525	6	E02573		E02573 DNA encodin
33	486.8	44.9	525	6	E07164		E07164 cDNA encodi
34	486.8	44.9	525	6	E15131		E15131 Human mRNA
35	486.8	44.9	525	6	I71150		I71150 Sequence 1
36	486.8	44.9	525	6	I83709		I83709 Sequence 1
37	486.8	44.9	525	22	E09431		E09431 cDNA encodi
38	486.8	44.9	660	9	HUMGCSF		M17706 Human granu
39	486.8	44.9	1498	9	HSGCSFR1		X03655 Human mRNA
40	486.8	44.9	1521	6	E01219		E01219 cDNA encodi
41	486.8	44.9	1525	6	I05333		I05333 Sequence 11
42	486.8	44.9	1525	6	I08669		I08669 Sequence 1
43	486.8	44.9	1525	6	I09205		I09205 Sequence 5
44	485.2	44.8	522	6	E01631		E01631 DNA encodin
45	485.2	44.8	522	6	E08529		E08529 DNA encodin

ALIGNMENTS

RESULT 1

E11965 standard; RNA; HUM; 555 BP.

XX E11965;
AC E11965;
XX E11965.1

XX 08-OCT-1997 (Rel. 52, Created)
DT 02-SEP-2000 (Rel. 65, Last updated, Version 2)

XX Human cDNA encoding thrombopoietin.

XX JP 1996228781-A/6.

XX Homo sapiens (human)

OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.

XX [1]
RN

1-555
 RA Miyazaki H., Kato H., Ogami K., Iwamatsu A., Akahori H., Kuroki R.,
 RA Shimizu T., Muto T.;
 RT DNA CODING FOR PROTEIN HAVING TPO ACTIVITY*;
 RL Patent number JP1996228781-A/6, 10-SEP-1996.
 RL KIRIN BREWERY CO LTD.
 XX
 CC OS Homo sapiens (human)
 CC PN JP 1996228781-A/6
 CC PD 10-SEP-1996
 CC PR 14-FEB-1995 JP 1995063298
 CC PR 14-FEB-1994 JP 94P 39090, 25-MAR-1994 JP 94P -79842,
 CC PR 01-JUN-1994 JP 94P 155126, 15-JUN-1994 JP 94P 167328,
 CC PR 17-AUG-1994 JP 94P 227159, 17-AUG-1994 JP 94P 193169,
 CC PR 01-NOV-1994 JP 94P 304167, 01-DEC-1994 JP 94P 298669,
 CC PR 28-DEC-1994 JP 94P 341200
 CC PI MIYAZAKI HIROSHI, KATO HISASHI, OGAMI KINYA, IWAMATSU AKIHIKO,
 CC PI AKAHORI HITONORI, KUROKI RYOTA, SHIMIZU TOSHIYUKI,
 CC PI MUTO TAKANORI
 CC PC C12N15/09,A61K38/00,C07K14/52,C12N1/21,C12N5/10,C12P21/02,
 CC PC (C12N1/21),
 CC PC C12R1:19), (C12N5/10,C12R1:91), (C12P21/02,C12R1:19), (C12P21/02,
 CC PC C12R1:91);
 CC CC strandedness: Double;
 CC CC topology: Linear;
 CC FH key Location/Qualifiers
 CC FH source 1..555
 CC FT /organism="Homo sapiens"
 CC FT /tissue_type="liver",
 CC FT CDS 1..555
 CC FT /product="thrombopoietin"
 CC FT sig_peptide 1..63
 XX
 FH Key Location/Qualifiers
 FH source 1..555
 FH /db_xref="taxon:9606"
 FH /organism="Homo sapiens"
 SQ Sequence 555 BP; 106 A; 183 C; 145 G; 121 T; 0 other;

Query Match 48.3%; Score 523.2; DB 22; Length 555;
 Best Local Similarity 99.4%; Pred. No. 5.9e-99;
 Matches 525; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 ATGGAGCTGACTGAATTCCTCGTGGTCAATGCTTCTCTTAAGTCAAGGCTAACGCTG 60
 DB 1 ATGGAGCTGACTGAATTCCTCGTGGTCAATGCTTCTCTTAAGTCAAGGCTAACGCTG 60
 QY 61 TCCAGCCCGGCTCTCTCTGCTTGTGACCTCCGAGTCCAGTCCAGTAACTGCTTCGTAAGTCC 120
 DB 61 TCCAGCCCGGCTCTCTCTGCTTGTGACCTCCGAGTCCAGTAACTGCTTCGTAAGTCC 120
 QY 121 CATGTCTCTTACAGCAGACTGAGCCAGTCCAGAGTCCAGAGTTCACCTTTGCTTACACCTGTC 180
 DB 121 CATGTCTCTTACAGCAGACTGAGCCAGTCCAGAGTTCACCTTTGCTTACACCTGTC 180
 QY 181 CTGCTCCCTGCTGTGGACTTTAGCTTGGAGAAATGGAAACCCAGATGGAGAGACCAAG 240
 DB 181 CTGCTCCCTGCTGTGGACTTTAGCTTGGAGAAATGGAAACCCAGATGGAGAGACCAAG 240
 QY 241 GCACAGGACATTTCTGGAGCAGTGAACCTTCTGCTGGAGGAGTGTGCGACACGGGA 300
 DB 241 GCACAGGACATTTCTGGAGCAGTGAACCTTCTGCTGGAGGAGTGTGCGACACGGGA 300
 QY 301 CAACCTGGGACCCACTTGCCTCTCATCCCTCTGCTGGAGGAGTGTGCGACACGGGA 360
 DB 301 CAACCTGGGACCCACTTGCCTCTCATCCCTCTGCTGGAGGAGTGTGCGACACGGGA 360
 QY 361 CTCCTTGGGGCCCTGCAGAGCCTCTTGGAAACCCAGCTTCTCCACAGGGCAGGACACCA 420
 DB 361 CTCCTTGGGGCCCTGCAGAGCCTCTTGGAAACCCAGCTTCTCCACAGGGCAGGACACCA 420
 QY 421 GCTCACAAGGATCCCAATGCCATCTTCTGAGCTTCCACACCTGCTCCGAGGAAAGGTG 480
 DB 421 GCTCACAAGGATCCCAATGCCATCTTCTGAGCTTCCACACCTGCTCCGAGGAAAGGTG 480
 QY 481 CGTTTCTCTGATGCTGTAGAGGGTCCACCTCTGCTGCTGAGGGTGGC 528

Db 361 CTCCTTGGGGCCCTGCAGAGCCTCTTGGAAACCCAGCTTCTCCACAGGGCAGGACCA 420
 QY 421 GCTCACAAGGATCCCAATGCCATCTTCTGAGCTTCCACACCTGCTCCGAGGAAAGGTG 480
 Db 421 GCTCACAAGGATCCCAATGCCATCTTCTGAGCTTCCACACCTGCTCCGAGGAAAGGTG 480
 QY 481 CGTTTCTCTGATGCTGTAGAGGGTCCACCTCTGCTGCTGAGGGTGGC 528
 Db 481 CGTTTCTCTGATGCTGTAGAGGGTCCACCTCTGCTGCTGAGGGTGGC 528
 RESULT 2
 ARI03959
 LOCUS AR103959 1059 bp DNA PAT 14-FEB-2001
 DEFINITION Sequence 57 from patent US 6093395.
 ACCESSION AR103959
 VERSION AR103959.1 GI:12816667
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 1059)
 AUTHORS Bauer S Christopher, Abrams M, Allen, Braford-Goldberg S, Ruth, Caparon M, Helena, Easton A, Michael, Klein B, Kure, McKeown J, P., Olin, P. O., Paik K, and Thomas J, W.
 TITLE Co-administration of interleukin-3 mutant polypeptides with CSF's for multi-lineage hematopoietic cell production
 JOURNAL Patent: US 6093395-A 57 25-JUL-2000;
 FEATURES Location/Qualifiers
 source 1..1059
 BASE COUNT 221 a 375 c 236 g 227 t
 ORIGIN
 Query Match 48.3%; Score 523.2; DB 6; Length 1059;
 Best Local Similarity 99.4%; Pred. No. 5e-99;
 Matches 525; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 ATGGAGCTGACTGAATTCCTCGTGGTCAATGCTTCTCTTAAGTCAAGGCTAACGCTG 60
 DB 1 ATGGAGCTGACTGAATTCCTCGTGGTCAATGCTTCTCTTAAGTCAAGGCTAACGCTG 60
 QY 61 TCCAGCCCGGCTCTCTCTGCTTGTGACCTCCGAGTCCAGTAACTGCTTCGTAAGTCC 120
 DB 61 TCCAGCCCGGCTCTCTCTGCTTGTGACCTCCGAGTCCAGTAACTGCTTCGTAAGTCC 120
 QY 121 CATGTCTCTTACAGCAGACTGAGCCAGTCCAGAGTCCAGAGTTCACCTTTGCTTACACCTGTC 180
 DB 121 CATGTCTCTTACAGCAGACTGAGCCAGTCCAGAGTTCACCTTTGCTTACACCTGTC 180
 QY 181 CTGCTCCCTGCTGTGGACTTTAGCTTGGAGAAATGGAAACCCAGATGGAGAGACCAAG 240
 DB 181 CTGCTCCCTGCTGTGGACTTTAGCTTGGAGAAATGGAAACCCAGATGGAGAGACCAAG 240
 QY 241 GCACAGGACATTTCTGGAGCAGTGAACCTTCTGCTGGAGGAGTGTGCGACACGGGA 300
 DB 241 GCACAGGACATTTCTGGAGCAGTGAACCTTCTGCTGGAGGAGTGTGCGACACGGGA 300
 QY 301 CAACCTGGGACCCACTTGCCTCTCATCCCTCTGCTGGAGGAGTGTGCGACACGGGA 360
 DB 301 CAACCTGGGACCCACTTGCCTCTCATCCCTCTGCTGGAGGAGTGTGCGACACGGGA 360
 QY 361 CTCCTTGGGGCCCTGCAGAGCCTCTTGGAAACCCAGCTTCTCCACAGGGCAGGACACCA 420
 DB 361 CTCCTTGGGGCCCTGCAGAGCCTCTTGGAAACCCAGCTTCTCCACAGGGCAGGACACCA 420
 QY 421 GCTCACAAGGATCCCAATGCCATCTTCTGAGCTTCCACACCTGCTCCGAGGAAAGGTG 480
 DB 421 GCTCACAAGGATCCCAATGCCATCTTCTGAGCTTCCACACCTGCTCCGAGGAAAGGTG 480
 QY 481 CGTTTCTCTGATGCTGTAGAGGGTCCACCTCTGCTGCTGAGGGTGGC 528

Db 481 CGTTTCCTGATGCTTGTAGGAGGTCCACCTCTGCCTCAGCGGGCC 528

RESULT 3

LOCUS AR087133 1062 bp DNA PAT 07-SEP-2000
DEFINITION Sequence 3 from patent US 5986049.
ACCESSION AR087133
VERSION AR087133.1 GI:10013899
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.

REFERENCE 1 (bases 1 to 1062)

AUTHORS Forstrom,J.W., Lofton-Day,C.E. and Lok,S.
TITLE Purified thrombopoietin and method of making it
JOURNAL Patent: US 5986049-A 3 16-NOV-1999;
FEATURES Location/Qualifiers
source 1..1062
BASE COUNT 223 a 375 c 236 g 228 t
ORIGIN

Query Match 48.3%; Score 523.2; DB 6; Length 1062;

Best Local Similarity 99.4%; Pred. No. 5e-99; Mismatches 0; Gaps 0;

QY 1 ATGGAGCTGACTGAATTGCTCTCGTGGTCAATGCTTCTCTTAAGCTGAAGGCTAACGCTG 60
|||||
Db 1 ATGGAGCTGACTGAATTGCTCTCGTGGTCAATGCTTCTCTTAAGCTGAAGGCTAACGCTG 60
|||||
QY 61 TCCAGCCGGCTCCTCGTCTGTTGACCTCGGAGTCTCAGTAACTGCTTCGTAAGTCC 120
|||||
Db 61 TCCAGCCGGCTCCTCGTCTGTTGACCTCGGAGTCTCAGTAACTGCTTCGTAAGTCC 120
|||||
QY 121 CATGCTCTTACAGCAGACTGAGCCAGTGCCTGAGGAGTTCACCCCTTTGCCTACACCTGTC 180
|||||
Db 121 CATGCTCTTACAGCAGACTGAGCCAGTGCCTGAGGAGTTCACCCCTTTGCCTACACCTGTC 180
|||||
QY 181 CTGCTGCTGCTGAGACTTTAGCTTGGGAGATGGAACCCAGATGGAGGAGGACCAAG 240
|||||
Db 181 CTGCTGCTGCTGAGACTTTAGCTTGGGAGATGGAACCCAGATGGAGGAGGACCAAG 240
|||||
QY 241 GCACAGGACATTCCTGGGAGCAGTCAACCTTCTGCTGGAGGAGTGGGAGGAGGAGG 300
|||||
Db 241 GCACAGGACATTCCTGGGAGCAGTCAACCTTCTGCTGGAGGAGTGGGAGGAGGAGG 300
|||||
QY 301 CAACTGGGACCCACTTGCCTCTCATCCCTCCTGGGGCAGCTTTCTGGACAGGTCGGTCTC 360
|||||
Db 301 CAACTGGGACCCACTTGCCTCTCATCCCTCCTGGGGCAGCTTTCTGGACAGGTCGGTCTC 360
|||||
QY 361 CTCCTTGGGGCCCTGCGAGAGCTCCTTGGAAACCCAGCTTCTCCACAGGGCAGGACCA 420
|||||
Db 361 CTCCTTGGGGCCCTGCGAGAGCTCCTTGGAAACCCAGCTTCTCCACAGGGCAGGACCA 420
|||||
QY 421 GCTCACAAGGATCCCAATGCCATCTTCTGAGCTTCCAAACACCTGCTCCGAGGAAAGTG 480
|||||
Db 421 GCTCACAAGGATCCCAATGCCATCTTCTGAGCTTCCAAACACCTGCTCCGAGGAAAGTG 480
|||||
QY 481 CGTTTCCTGATGCTTGTAGGAGGTCCACCTCTGCGTCAAGGGTGGC 528
|||||
Db 481 CGTTTCCTGATGCTTGTAGGAGGTCCACCTCTGCGTCAAGGGTGGC 528
|||||

RESULT 4

LOCUS AR087571 1062 bp DNA PAT 07-SEP-2000
DEFINITION Sequence 18 from patent US 5989537.
ACCESSION AR087571
VERSION AR087571.1 GI:10014334
KEYWORDS
SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 1062)

AUTHORS Holly,R.D., Lok,S., Foster,D.C., Hagen,F.S., Kaushansky,K.,
Kuijper,J.L., Lofton-Day,C.E. and Oort,P.J.
TITLE Methods for stimulating granulocyte/macrophage lineage using
thrombopoietin
JOURNAL Patent: US 5989537-A 18 23-NOV-1999;
FEATURES Location/Qualifiers
source 1..1062
BASE COUNT 223 a 375 c 236 g 228 t
ORIGIN

Query Match 48.3%; Score 523.2; DB 6; Length 1062;
Best Local Similarity 99.4%; Pred. No. 5e-99; Mismatches 0; Gaps 0;

QY 1 ATGGAGCTGACTGAATTGCTCTCGTGGTCAATGCTTCTCTTAAGCTGAAGGCTAACGCTG 60
|||||
Db 1 ATGGAGCTGACTGAATTGCTCTCGTGGTCAATGCTTCTCTTAAGCTGAAGGCTAACGCTG 60
|||||
QY 61 TCCAGCCGGCTCCTCGTCTGTTGACCTCGGAGTCTCAGTAACTGCTTCGTAAGTCC 120
|||||
Db 61 TCCAGCCGGCTCCTCGTCTGTTGACCTCGGAGTCTCAGTAACTGCTTCGTAAGTCC 120
|||||
QY 121 CATGCTCTTACAGCAGACTGAGCCAGTGCCTGAGGAGTTCACCCCTTTGCCTACACCTGTC 180
|||||
Db 121 CATGCTCTTACAGCAGACTGAGCCAGTGCCTGAGGAGTTCACCCCTTTGCCTACACCTGTC 180
|||||
QY 181 CTGCTGCTGCTGAGACTTTAGCTTGGGAGATGGAACCCAGATGGAGGAGGACCAAG 240
|||||
Db 181 CTGCTGCTGCTGAGACTTTAGCTTGGGAGATGGAACCCAGATGGAGGAGGACCAAG 240
|||||
QY 241 GCACAGGACATTCCTGGGAGCAGTCAACCTTCTGCTGGAGGAGTGGGAGGAGGAGG 300
|||||
Db 241 GCACAGGACATTCCTGGGAGCAGTCAACCTTCTGCTGGAGGAGTGGGAGGAGGAGG 300
|||||
QY 301 CAACTGGGACCCACTTGCCTCTCATCCCTCCTGGGGCAGCTTTCTGGACAGGTCGGTCTC 360
|||||
Db 301 CAACTGGGACCCACTTGCCTCTCATCCCTCCTGGGGCAGCTTTCTGGACAGGTCGGTCTC 360
|||||
QY 361 CTCCTTGGGGCCCTGCGAGAGCTCCTTGGAAACCCAGCTTCTCCACAGGGCAGGACCA 420
|||||
Db 361 CTCCTTGGGGCCCTGCGAGAGCTCCTTGGAAACCCAGCTTCTCCACAGGGCAGGACCA 420
|||||
QY 421 GCTCACAAGGATCCCAATGCCATCTTCTGAGCTTCCAAACACCTGCTCCGAGGAAAGTG 480
|||||
Db 421 GCTCACAAGGATCCCAATGCCATCTTCTGAGCTTCCAAACACCTGCTCCGAGGAAAGTG 480
|||||
QY 481 CGTTTCCTGATGCTTGTAGGAGGTCCACCTCTGCGTCAAGGGTGGC 528
|||||
Db 481 CGTTTCCTGATGCTTGTAGGAGGTCCACCTCTGCGTCAAGGGTGGC 528
|||||

RESULT 5

LOCUS I49760 1062 bp DNA PAT 07-OCT-1997
DEFINITION Sequence 3 from patent US 5641655.
ACCESSION I49760
VERSION I49760.1 GI:2471980
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.

REFERENCE 1 (bases 1 to 1062)

AUTHORS Foster,D.C., Heipel,M.D. and Holly,R.D.
TITLE Methods for producing thrombopoietin polypeptides using a mammalian
tissue plasminogen activator secretory peptide
JOURNAL Patent: US 5641655-A 3 24-JUN-1997;
FEATURES Location/Qualifiers
source 1..1062

COMMENT	OS	PN	JP	PN	JP
	KIRIN BREWERY CO LTD	OS Homo sapiens (human)	PN JP 1998212243-A/3		

15.

DEFINITION	Human cDNA encoding thrombopoietin.
ACCESSION	E12182
VERSION	E12182.1 GI:3251016
KEYWORDS	JP 1996277296-A/2.
SOURCE	Homo sapiens.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Crani-
REFERENCE	Mammalia; Eutheria; Primates; Catari-
REFERENCE	1 (bases 1 to 1267)
REFERENCE	Miyazaki,H., Kato,H., Ogami,K., Iwa-
REFERENCE	Shimizu,T. and Muto,T.
TITLE	PROTEIN HAVING RPO ACTIVITY
JOURNAL	Patent: JP 1996277296-A 2 22-OCT-1996
COMMENT	KIRIN BREWERY CO LTD
	OS Homo sapiens (human)
	PN JP 1996277296-A/2
	PD 22-OCT-1996
	PF 14-FEB-1995 JP 1995161363
	PR 14-FEB-1994 JP 94P 39090, 2
	01-JUN-1994 JP 94P 155126, 15-JUN-
	17-AUG-1994 JP 94P 227159, 17-DEC-
	01-NOV-1994 JP 94P 304167, 01-AUG-
	28-DEC-1994 JP 94P 341200
	PI MIYAZAKI HIROSHI, KATO HISASHI
	PI AKAHORI HIRONORI, KUROKI RYOTARU
	TAKANORI
	PC C07K14/52.C07H21/04.C12N15/09.

```

CC (c1azf21/02,c1azn1.51),
CC strandedness: Double;
CC topology: Linear;
FH key Location/Qualifiers
FH
FT source 1. .1267
FT /organism='Homo sapiens'
FT /tissue_type='liver'
FT CDS 25. .1086
FT /product='thrombopoietin'
FT sig_peptide 25. .87.
FT Location/Qualifiers
FT 1. .1267
FT /organism='Homo sapiens'
FT /db_xref='taxon:9606'
BASE COUNT 281 a 436 c 279 g 271 t
ORIGIN
Query Match 48.3%; Score 523.2; DB 6; Length 1267;
Best Local Similarity 99.4%; Pred. No. 4.8e-99;
Matches 525: Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

DB	263	GCACAGGACATCTGGGAGCAGTACCCCTTCTGCTGGAGGGAGTGAATGGCAGCACACGGGGA	324
QY	301	CAACTGGGACCACTTGCCTCTCATCCCTCCTGGGCAGCTTTCTGGACAGTCCGCTCTC	360

```

Db 325 CAACCTGGAGCCCACTTGCCTCTCATCCCTCCCTGGGGCAGCTTTCTGGACAGGTCCCTCTC 384
QY 361 CTCCTTGGGGCCCTCAGAGCCCTCTTGGAAACCCAGCTTCTCCACAGGGCAGGACCA 420
Db 385 CTCCTTGGGGCCCTCAGAGCCCTCTTGGAAACCCAGCTTCTCCACAGGGCAGGACCA 444
QY 421 GCTCACAGGATCCCAATGCTTCCTGAGCTTCCACACCTGCTCCGAGGAAAGGTG 480
Db 445 GCTCACAGGATCCCAATGCTTCCTGAGCTTCCACACCTGCTCCGAGGAAAGGTG 504
QY 481 CGTTCTGATGCTTGTAGGAGGCTCCACCCCTGCGTCCAGGGTGGC 528
Db 505 CGTTCTGATGCTTGTAGGAGGCTCCACCCCTGCGTCCAGGGTGGC 552

RESULT 9
E12214
LOCUS E12214 1267 bp DNA PAT 24-JUN-1998
DEFINITION Human cDNA encoding thrombopoietin.
ACCESSION E12214
VERSION E12214.1 GI:3251048
KEYWORDS JP 1996291196-A/1.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1267)
AUTHORS Miyazaki,H., Kato,H., Ogami,K., Iwamatsu,A., Akahori,H., Kuroki,R.,
Shinizu,T. and Muto,T.
TITLE PROTEIN HAVING TPO ACTIVITY.
JOURNAL Patent: JP 1996291196-A 1 05-NOV-1996;
KIRIN BREWERY CO LTD
COMMENT OS Homo sapiens (human)
PN JP 1996291196-A/1
PD 05-NOV-1996
PF 14-FEB-1995 JP 1995355052
PR 14-FEB-1994 JP 94P 39090, 25-MAR-1994 JP 94P 79842, PR
01-JUN-1994 JP 94P 155126, 15-JUN-1994 JP 94P 167328, PR
17-AUG-1994 JP 94P 227159, 17-AUG-1994 JP 94P 193169, PR
01-NOV-1994 JP 94P 304167, 01-DEC-1994 JP 94P 298669, PR
28-DEC-1994 JP 94P 341200
PI MIYAZAKI HIROSHI, KATO HISASHI, OGAMI KINYA, IWAMATSU AKIHIKO,
PI AKAHORI HIRONORI, KUROKI RYOTA, SHIMIZU TOSHIYUKI, PI MUTO
TAKANORI
PC C07K14/53,C07K1/22,C12N15/09//A61K38/00,C12P21/08; CC
strandedness: Double;
CC topology: Linear;
FH Key Location/Qualifiers
FT source 1..1267
FT /organism='Homo sapiens'
FT /tissue_type='liver'
FT CDS 25..1086
FT /product='thrombopoietin'
FT sig_peptide 25..87
FT mat_peptide 88..1083
FT /product='thrombopoietin'.
FEATURES
source 1..1267
/organism='Homo sapiens'
/db_xref='taxon:9606'
BASE COUNT 281 a 436 c 279 g 271 t
ORIGIN

Query Match 48.3%; Score 523.2; DB 6; Length 1267;
Best Local Similarity 99.4%; Pred. No. 4.8e-99;
Matches 525; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGGAGCTGACTGAATTGCTCTGCTGGTCAATGCTTCTCCTAACTGCAAGGCTAACGCTG 60
Db 25 ATGGAGCTGACTGAATTGCTCTGCTGGTCAATGCTTCTCCTAACTGCAAGGCTAACGCTG 84

```

```

QY 61 TCCAGCCCCGGCTCCCTGCTCTGCTTGTGACCTCCGAGTCCCTCAGTAACTGCTTCTGACTCC 120
Db 85 TCCAGCCCCGGCTCCCTGCTCTGCTTGTGACCTCCGAGTCCCTCAGTAACTGCTTCTGACTCC 144
QY 121 CATGTCTTCACAGCAGACTGAGCCAGTCCCAGAGAGGTTTCAACCTTTGCCCTACACCTGTC 180
Db 145 CATGTCTTCACAGCAGACTGAGCCAGTCCCAGAGAGGTTTCAACCTTTGCCCTACACCTGTC 204
QY 181 CTGTCTGCCCTGCTGTGAGCTTTAGCTTTGGGAGAAATGGAAACCCAGATGGAGGAGACCAAG 240
Db 205 CTGTCTGCCCTGCTGTGAGCTTTAGCTTTGGGAGAAATGGAAACCCAGATGGAGGAGACCAAG 264
QY 241 GCACAGGACATTTCTGGGAGCAGTACCCTTCTGCTGGAGGGAGTGTATGGCAGCACGGGGA 300
Db 265 GCACAGGACATTTCTGGGAGCAGTACCCTTCTGCTGGAGGGAGTGTATGGCAGCACGGGGA 324
QY 301 CAACTGGGACCCACTTGCCTCTCATCCCTCCCTGGGGCAGCTTTCTTGACAGAGTCCGCTCTC 360
Db 325 CAACTGGGACCCACTTGCCTCTCATCCCTCCCTGGGGCAGCTTTCTTGACAGAGTCCGCTCTC 384
QY 361 CTCTTTGGGGCCCTGCAGAGCCTCCTTTGGAACCCAGCTTCTCCACAGGGCAGGACCA 420
Db 385 CTCTTTGGGGCCCTGCAGAGCCTCCTTTGGAACCCAGCTTCTCCACAGGGCAGGACCA 444
QY 421 GCTCAAGGATCCCAATGCCATCTTCTGAGCTTCCAAACACCTGCTCCGAGGAAAGGTG 480
Db 445 GCTCAAGGATCCCAATGCCATCTTCTGAGCTTCCAAACACCTGCTCCGAGGAAAGGTG 504
QY 481 CGTTTCTGATGCTTGTAGGAGGGTCCACCCCTGCGTCAGGGTGGC 528
Db 505 CGTTTCTGATGCTTGTAGGAGGGTCCACCCCTGCGTCAGGGTGGC 552

RESULT 10
AR008878
LOCUS AR008878 1342 bp DNA PAT 04-DEC-1998
DEFINITION Sequence 1 from patent US 5756083.
ACCESSION AR008878
VERSION AR008878.1 GI:3967683
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1342)
AUTHORS Elliott,S.G.
TITLE Mpl ligand analogs
JOURNAL Patent: US 5756083-A 1 26-MAY-1998;
FEATURES Location/Qualifiers
source 1..1342
/organism='unknown'
BASE COUNT 299 a 454 c 293 g 296 t
ORIGIN

Query Match 48.3%; Score 523.2; DB 6; Length 1342;
Best Local Similarity 99.4%; Pred. No. 4.7e-99;
Matches 525; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGGAGCTGACTGAATTGCTCTGCTGGTCAATGCTTCTCCTAACTGCAAGGCTAACGCTG 60
Db 36 ATGGAGCTGACTGAATTGCTCTGCTGGTCAATGCTTCTCCTAACTGCAAGGCTAACGCTG 95
QY 61 TCCAGCCCCGGCTCCCTGCTCTGCTTGTGACCTCCGAGTCCCTCAGTAACTGCTTCTGACTCC 120
Db 96 TCCAGCCCCGGCTCCCTGCTCTGCTTGTGACCTCCGAGTCCCTCAGTAACTGCTTCTGACTCC 155
QY 121 CATGTCTTCACAGCAGACTGAGCCAGTCCCAGAGAGTTCACCCCTTTGCCCTACACCTGTC 180
Db 156 CATGTCTTCACAGCAGACTGAGCCAGTCCCAGAGAGTTCACCCCTTTGCCCTACACCTGTC 215
QY 181 CTGTGCTCTGCTGTGAGACTTTAGCTTTGGGAGAAATGGAAACCCAGATGGAGGAGACCAAG 240
Db 216 CTGTGCTCTGCTGTGAGACTTTAGCTTTGGGAGAAATGGAAACCCAGATGGAGGAGACCAAG 275

```


QY 241 GCACAGACATTCGGGAGCAGTACCTTCTGCTGGAGGAGTGGCAGCAGCGGA 300
 Db 276 GCACAGACATTCGGGAGCAGTACCTTCTGCTGGAGGAGTGGCAGCAGCGGA 335
 QY 301 CAACTGGAGCCACTTCCCTCTCATCCCTCTCTGGGGAGCTTTCTGGACAGGTCCTCTC 360
 Db 336 CAACTGGAGCCACTTCCCTCTCATCCCTCTCTGGGGAGCTTTCTGGACAGGTCCTCTC 395
 QY 361 CTCTTGGGCGCTTCAGAGCCCTTGGAAACCAGCTTCTCCACAGGCGAGACACA 420
 Db 396 CTCTTGGGCGCTTCAGAGCCCTTGGAAACCAGCTTCTCCACAGGCGAGACACA 455
 QY 421 GCTCACAGGATCCCAATGCCATCTTCTGAGCTTCAACACCTTCCAGGAAAGGTG 480
 Db 456 GCTCACAGGATCCCAATGCCATCTTCTGAGCTTCAACACCTTCCAGGAAAGGTG 515
 QY 481 CGTTTCTGTATGCTTGTAGGAGGTCACCCCTCTGCGTCAGGGGTGGC 528
 Db 516 CGTTTCTGTATGCTTGTAGGAGGTCACCCCTCTGCGTCAGGGGTGGC 563

RESULT 11

AR023468
 LOCUS AR023468 1342 bp DNA PAT 05-DEC-1998
 DEFINITION Sequence 24 from patent US 5795569.
 ACCESSION AR023468
 VERSION AR023468.1 GI:3976762
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 1342)
 AUTHORS Bartley T.D., Bogenberger, J.M., Bosselman, R.A., Hunt, P.,
 Kinstler, O.B. and Samal, B.B.
 TITLE Mono-pegylated proteins that stimulate megakaryocyte growth and
 differentiation
 JOURNAL Patent: US 5795569-A 24 18-AUG-1998;
 FEATURES Location/Qualifiers
 source 1..1342
 BASE COUNT 299 a 454 c 293 g 296 t
 ORIGIN

Query Match 48.3%; Score 523.2; DB 6; Length 1342;
 Best Local Similarity 99.4%; Pred. No. 4.7e-99;
 Matches 525; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGAGCTGACTGAATTCCTCTCGTGGTATGCTTCTTAAGTGAAGGCTAACGCTG 60
 Db 36 ATGAGCTGACTGAATTCCTCTCGTGGTATGCTTCTTAAGTGAAGGCTAACGCTG 95
 QY 61 TCAGCCCGGCTCCTCTGCTGTGACCTCCGAGTCCCTAGTAACTGCTTCGTGACTCC 120
 Db 96 TCAGCCCGGCTCCTCTGCTGTGACCTCCGAGTCCCTAGTAACTGCTTCGTGACTCC 155
 QY 121 CATGCTCTTACAGCAGACTGACCCAGTCCGAGAGTTTACCCCTTGCCTACACCTGTC 180
 Db 156 CATGCTCTTACAGCAGACTGACCCAGTCCGAGAGTTTACCCCTTGCCTACACCTGTC 215
 QY 181 CTGCTGCTCTGAGCTTTAGCTTGGGAGATGGAACCCAGATGGAGGAGACCAAG 240
 Db 216 CTGCTGCTCTGAGCTTTAGCTTGGGAGATGGAACCCAGATGGAGGAGACCAAG 275
 QY 241 GCACAGACATTCCTGGAGCAGTGACCTTCTGCTGGAGGAGTGTATGGCAGCAGCGGA 300
 Db 276 GCACAGACATTCCTGGAGCAGTGACCTTCTGCTGGAGGAGTGTATGGCAGCAGCGGA 335
 QY 301 CAACTGGAGCCACTTCCCTCTCATCCCTCTGGGGAGCTTTCTGGACAGGTCCTCTC 360
 Db 336 CAACTGGAGCCACTTCCCTCTCATCCCTCTGGGGAGCTTTCTGGACAGGTCCTCTC 395

QY 361 CTCTTGGGCGCTTCAGAGCCCTTGGAAACCAGCTTCTCCACAGGCGAGACACA 420
 Db 396 CTCTTGGGCGCTTCAGAGCCCTTGGAAACCAGCTTCTCCACAGGCGAGACACA 455
 QY 421 GCTCACAGGATCCCAATGCCATCTTCTGAGCTTCAACACCTTCCAGGAAAGGTG 480
 Db 456 GCTCACAGGATCCCAATGCCATCTTCTGAGCTTCAACACCTTCCAGGAAAGGTG 515
 QY 481 CGTTTCTGTATGCTTGTAGGAGGTCACCCCTCTGCGTCAGGGGTGGC 528
 Db 516 CGTTTCTGTATGCTTGTAGGAGGTCACCCCTCTGCGTCAGGGGTGGC 563

RESULT 12

AR087573
 LOCUS AR087573 1342 bp DNA PAT 07-SEP-2000
 DEFINITION Sequence 1 from patent US 5989538.
 ACCESSION AR087573
 VERSION AR087573.1 GI:10014336
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 1342)
 AUTHORS Elliott, S.G.
 TITLE Mpl ligand analogs
 JOURNAL Patent: US 5989538-A 1 23-NOV-1999;
 FEATURES Location/Qualifiers
 source 1..1342
 BASE COUNT 299 a 454 c 293 g 296 t
 ORIGIN

Query Match 48.3%; Score 523.2; DB 6; Length 1342;
 Best Local Similarity 99.4%; Pred. No. 4.7e-99;
 Matches 525; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGAGCTGACTGAATTCCTCTCGTGGTATGCTTCTTAAGTGAAGGCTAACGCTG 60
 Db 36 ATGAGCTGACTGAATTCCTCTCGTGGTATGCTTCTTAAGTGAAGGCTAACGCTG 95
 QY 61 TCAGCCCGGCTCCTCTGCTGTGACCTCCGAGTCCCTAGTAACTGCTTCGTGACTCC 120
 Db 96 TCAGCCCGGCTCCTCTGCTGTGACCTCCGAGTCCCTAGTAACTGCTTCGTGACTCC 155
 QY 121 CATGCTCTTACAGCAGACTGACCCAGTCCGAGAGTTTACCCCTTGCCTACACCTGTC 180
 Db 156 CATGCTCTTACAGCAGACTGACCCAGTCCGAGAGTTTACCCCTTGCCTACACCTGTC 215
 QY 181 CTGCTGCTCTGAGCTTTAGCTTGGGAGATGGAACCCAGATGGAGGAGACCAAG 240
 Db 216 CTGCTGCTCTGAGCTTTAGCTTGGGAGATGGAACCCAGATGGAGGAGACCAAG 275
 QY 241 GCACAGACATTCCTGGAGCAGTGACCTTCTGCTGGAGGAGTGTATGGCAGCAGCGGA 300
 Db 276 GCACAGACATTCCTGGAGCAGTGACCTTCTGCTGGAGGAGTGTATGGCAGCAGCGGA 335
 QY 301 CAACTGGAGCCACTTCCCTCTCATCCCTCTGGGGAGCTTTCTGGACAGGTCCTCTC 360
 Db 336 CAACTGGAGCCACTTCCCTCTCATCCCTCTGGGGAGCTTTCTGGACAGGTCCTCTC 395
 QY 361 CTCTTGGGCGCTTCAGAGCCCTTGGAAACCAGCTTCTCCACAGGCGAGACACA 420
 Db 396 CTCTTGGGCGCTTCAGAGCCCTTGGAAACCAGCTTCTCCACAGGCGAGACACA 455
 QY 421 GCTCACAGGATCCCAATGCCATCTTCTGAGCTTCAACACCTTCCAGGAAAGGTG 480
 Db 456 GCTCACAGGATCCCAATGCCATCTTCTGAGCTTCAACACCTTCCAGGAAAGGTG 515
 QY 481 CGTTTCTGTATGCTTGTAGGAGGTCACCCCTCTGCGTCAGGGGTGGC 528
 Db 516 CGTTTCTGTATGCTTGTAGGAGGTCACCCCTCTGCGTCAGGGGTGGC 563

RESULT 13
LOCUS I85129 1342 bp DNA PAT 04-APR-1998
DEFINITION Sequence 1 from patent US 5696250.
ACCESSION I85129
VERSION I85129.1 GI:3022649
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1342)
AUTHORS Elliott,S.G.
TITLE DNA encoding megakaryocyte growth and development factor analogs
JOURNAL Patent: US 5696250-A 1 09-DEC-1997;
FEATURES
source Location/Qualifiers
1..1342
/organism="unknown"
BASE COUNT 299 a 454 c 293 g 296 t
ORIGIN

Query Match 48.3%; Score 523.2; DB 6; Length 1342;
Best Local Similarity 99.4%; Pred. No. 4.7e-99;
Matches 525; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 ATGGAGCTGACTGAATTCCTCGTGGTCATGCTTCTCTAACTGCAAGGCTAAGCGTG 60
DB 36 ATGGAGCTGACTGAATTCCTCGTGGTCATGCTTCTCTAACTGCAAGGCTAAGCGTG 95
QY 61 TCCAGCCCGGCTCCTCTGCTTGTGACCTCCGAGTCCCTCAGTAACTGCTTGTGACTCC 120
DB 96 TCCAGCCCGGCTCCTCTGCTTGTGACCTCCGAGTCCCTCAGTAACTGCTTGTGACTCC 155
QY 121 CATGCTCTTACAGCAGACTGACCGAGTCCGAGAGGTTACCCCTTGGCTACACTGTC 180
DB 156 CATGCTCTTACAGCAGACTGACCGAGTCCGAGAGGTTACCCCTTGGCTACACTGTC 215
QY 181 CTGCTGCTCTGCTGACTTTAGCTTGGAGATGAAACCCAGATGAGGAGACCAAG 240
DB 216 CTGCTGCTCTGCTGACTTTAGCTTGGAGATGAAACCCAGATGAGGAGACCAAG 275
QY 241 GCACAGACATCTGGAGAGTACCCCTTCTGCTGGAGGAGTATGTCAGCAGCGGGA 300
DB 276 GCACAGACATCTGGAGAGTACCCCTTCTGCTGGAGGAGTATGTCAGCAGCGGGA 335
QY 301 CAACTGGAGCCACTTGCCTCTCATCCCTCGGGGAGCTTCTGGAGAGGTCCGCTC 360
DB 336 CAACTGGAGCCACTTGCCTCTCATCCCTCGGGGAGCTTCTGGAGAGGTCCGCTC 395
QY 361 CTCTTTGGGCGCTCGAGAGCTTCTTGGAAACCCAGCTTCTCCAGGGCAGGACCA 420
DB 396 CTCTTTGGGCGCTCGAGAGCTTCTTGGAAACCCAGCTTCTCCAGGGCAGGACCA 455
QY 421 GCTCACAAGATCCCAATGCAATCTTCTGAGCTTCCACACCTGCTCCGAGGAAAGTG 480
DB 456 GCTCACAAGATCCCAATGCAATCTTCTGAGCTTCCACACCTGCTCCGAGGAAAGTG 515
QY 481 CGTTTCTGATCTCTAGAGGGTCCACCCCTCGGTCAGGGTGGC 528
DB 516 CGTTTCTGATCTCTAGAGGGTCCACCCCTCGGTCAGGGGCGC 563

RESULT 14
LOCUS D32047 1691 bp mRNA PRI 03-SEP-1997
DEFINITION Homo sapiens mRNA for thrombopoietin, complete cds.
ACCESSION D32047
VERSION D32047.1 GI:2351117
KEYWORDS
SOURCE Homo sapiens liver cdna to mRNA.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Qami,K.
Direct Submission
Submitted (08-JUL-1994) to the DBJ/EMBL/GenBank databases. Kinya
Ogami, KIRIN Brewery Co., Ltd., Pharmaceutical Research Laboratory;
2-2 Soujamachi 1 chome, Maebashi, Gunma 371, Japan
(Tel:0272-54-8618, Fax:0272-52-2307)
2 (sites)
Kato,T., Ogami,K., Shimada,Y., Iwamatsu,A., Sohma,Y., Akahori,H.,
Horie,K., Kokubo,A., Kudo,Y., Maeda,E., Kobayashi,K., Ohashi,H.,
Ozawa,T., Inoue,H., Kawamura,K. and Miyazaki,H.
Purification and characterization of thrombopoietin
J. Biochem. 118 (1), 229-236 (1995)
96015174
Erratum: [[published erratum appears in J Biochem (Tokyo) 1996
Jan;119(1):208]]
FEATURES
Location/Qualifiers
1..1691
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="liver"
102..1163
102..1164
/codon_start=1
/product="thrombopoietin"
/protein_id="BAA21930.1"
/db_xref="GI:2351118"
/translation="MELTELLLVMLLTARLTLSPPACDLRLVSKLLRDSHVLIH
SRISQCEVHPILPTVLPVDFSLGEWKTOMEETKAODILGAVTLLLEGVMAARGQL
GPTCLSLGLGSLGQVRLLAGLALQLPQGTTHAKDPNATLFSQHLLRGKV
RFLMLVGSTLVCRRAPPTTAVPSRTSLVLTNLNPNRTSGLETTASARTGSGL
LKWOGFRKIPGLLNOTSRSDIOPGLNRLHELLNGRLGFLPGPSRRTLAGPDLS
GTSRGLSPNLPQGYSPSPHPPTGQVTLFPLPTLPVQLHLLPDPSPATPTP
TSPLLNTSYTHSONLSQEG"
BASE COUNT 383 a 542 c 371 g 395 t
ORIGIN

Query Match 48.3%; Score 523.2; DB 9; Length 1691;
Best Local Similarity 99.4%; Pred. No. 4.4e-99;
Matches 525; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 ATGGAGCTGACTGAATTCCTCGTGGTCATGCTTCTCTAACTGCAAGGCTAAGCGTG 60
DB 102 ATGGAGCTGACTGAATTCCTCGTGGTCATGCTTCTCTAACTGCAAGGCTAAGCGTG 161
QY 61 TCCAGCCCGGCTCCTCTGCTTGTGACCTCCGAGTCCCTCAGTAACTGCTTGTGACTCC 120
DB 162 TCCAGCCCGGCTCCTCTGCTTGTGACCTCCGAGTCCCTCAGTAACTGCTTGTGACTCC 221
QY 121 CATGCTCTTACAGCAGACTGACCGAGTCCGAGAGGTTACCCCTTGGCTACACTGTC 180
DB 222 CATGCTCTTACAGCAGACTGACCGAGTCCGAGAGGTTACCCCTTGGCTACACTGTC 281
QY 181 CTGCTGCTCTGCTGAGACTTTAGCTTGGAGATGAAACCCAGATGAGGAGACCAAG 240
DB 282 CTGCTGCTCTGCTGAGACTTTAGCTTGGAGATGAAACCCAGATGAGGAGACCAAG 341
QY 241 GCACAGGACATCTGGAGCAGTACCCCTTCTGCTGGAGGAGTATGTCAGCAGCGGGA 300
DB 342 GCACAGGACATCTGGAGCAGTACCCCTTCTGCTGGAGGAGTATGTCAGCAGCGGGA 401
QY 301 CAACTGGAGCCACTTGCCTCTCATCCCTCGGGGAGCTTCTGAGAGGTCCGCTC 360
DB 402 CAACTGGAGCCACTTGCCTCTCATCCCTCTCTGAGGAGCTTCTGAGAGGTCCGCTC 461
QY 361 CTCTTTGGGCGCTCGAGAGCTTCTTGGAAACCCAGCTTCTCCAGGGCAGGACCA 420
DB 462 CTCTTTGGGCGCTCGAGAGCTTCTTGGAAACCCAGCTTCTCCAGGGCAGGACCA 521
QY 421 GCTCACAAGATCCCAATGCAATCTTCTGAGCTTCCACACCTGCTCCGAGGAAAGTG 480

||||| 522 GCTCACAAGGATCCCAATGCCATCTCTCTAGCTTCCAAACACCTCTCGGAGGAAAGTG 581
QY 481 CGTTTCCCTGATGCTTGTAGGAGGGTCCACCTCTCGGTACGGGTGGC 528
Db 582 CGTTTCCCTGATGCTTGTAGGAGGGTCCACCTCTCGGTACGGGTGGC 629

RESULT 15
LOCUS E16668 1721 bp DNA PAT 28-JUL-1999
DEFINITION cDNA encoding thrombopoietin.
ACCESSION E16668
VERSION E16668.1 GI:5711351
KEYWORDS JP 1998212243-A/2.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1721)
Kuroda, K.
COMPOSITION CONTAINING THROMBOPOIETIN PROTEIN AND USED FOR
STIMULATING RECOVERY OF HEMOSTATIC THROMBUS-FORMING ABILITY OF
PATENT: JP 1998212243-A 2 11-AUG-1998;
KIRIN BREWERY CO LTD
OS Homo sapiens (human)
PN JP 1998212243-A/2
PD 11-AUG-1998
PF 31-JAN-1997 JP 1997019549
PI KURODA KENJI
PC A61K38/00,A01N1/02,C07K14/52//C12N15/09,C12P21/02,(C12P21/02,
C12R1:19);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
FH Key Location/Qualifiers
FT source 1..1721
FT /organism='Homo sapiens'
FT /tissue_type='Liver'
FT 1..101
FT 5'UTR
FT sig_peptide 102..164
FT /product='signal peptide of thrombopoietin' FT
FT mat_peptide 165..1160
FT /product='thrombopoietin'
FT CDS 102..1163
FT /product='thrombopoietin'
FT 1164..1721.
FT 3'UTR
FT Location/Qualifiers
source 1..1721
/organism='Homo sapiens'
/db_xref='taxon:9606'
BASE COUNT 413 a 542 c 371 g 395 t
ORIGIN

Query Match 48.3%; Score 523.2; DB 6; Length 1721;
Best Local Similarity 99.4%; Pred. No. 4.4e-99;
Matches 525; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGGAGCTGACTGAATTCCTCTCGTGTGTCATGCTTCTCTCACTGCAAGGCTAACGCTG 60
Db 102 ATGGAGCTGACTGAATTCCTCTCGTGTGTCATGCTTCTCTCACTGCAAGGCTAACGCTG 161
QY 61 TCCAGCCCGGCTCTCTCTGCTTGTGACCTCCGAGTCCAGTAAACTGCTTCGTCGACTCC 120
Db 162 TCCAGCCCGGCTCTCTCTGCTTGTGACCTCCGAGTCCAGTAAACTGCTTCGTCGACTCC 221
QY 121 CATGTCCTTACAGCAGACTGAGCCAGTGCCAGAGGTTACCCCTTTGCTTACACCTGTC 180
Db 222 CATGTCCTTACAGCAGACTGAGCCAGTGCCAGAGGTTACCCCTTTGCTTACACCTGTC 281

QY 181 CTGCTGCCTGCTGTGACTTTAGCTTGGAGAAATGAAAAACCCAGATGGAGGAGACCAAG 240
Db 282 CTGCTGCCTGCTGTGACTTTAGCTTGGAGAAATGAAAAACCCAGATGGAGGAGACCAAG 341
QY 241 GCACAGGACATTCCTGGGAGCAGTACCCTTCTGTGGAGGGAGTATGGCAGCAGCGGGA 300
Db 342 GCACAGGACATTCCTGGGAGCAGTACCCTTCTGTGGAGGGAGTATGGCAGCAGCGGGA 401
QY 301 CAACTGGGACCCACTTGCTCTCATCCCTCTGGGGCAGCTTCTGGACAGGTCCGTCTC 360
Db 402 CAACTGGGACCCACTTGCTCTCATCCCTCTGGGGCAGCTTCTGGACAGGTCCGTCTC 461
QY 361 CTCCTTGGGGCCCTGCAGAGCTCTTGGAAACCCAGCTTCTCCACAGGGCAGGACCA 420
Db 462 CTCCTTGGGGCCCTGCAGAGCTCTTGGAAACCCAGCTTCTCCACAGGGCAGGACCA 521
QY 421 GCTCACAAGGATCCCAATGCCATCTCTCTGAGCTTCCAAACACCTCTCCGAGGAAAGTG 480
Db 522 GCTCACAAGGATCCCAATGCCATCTCTCTGAGCTTCCAAACACCTCTCCGAGGAAAGTG 581
QY 481 CGTTTCCCTGATGCTTGTAGGAGGGTCCACCTCTCGGTACGGGTGGC 528
Db 582 CGTTTCCCTGATGCTTGTAGGAGGGTCCACCTCTCGGTACGGGTGGC 629

Search completed: April 5, 2002, 10:29:34
Job time: 4767 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 5, 2002, 10:33:05 ; Search time 188.75 Seconds
(without alignments)
4919.116 Million cell updates/sec

Title: US-09-680-514-6
Perfect score: 1083
Sequence: 1 ATGGAGCTGACTGAATTGCT.....TACGCCACCTTGCCAGCC 1083

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues
Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_1101.*
1: /SIDS2/gcgdata/geneseq/geneseq/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1083	100.0	1083	17	AAT41787
2	1030.2	95.1	1095	17	AAT41788
3	993	91.7	1047	17	AAT41786
4	523.2	48.3	861	16	AAQ99552
5	523.2	48.3	1062	16	AAQ03941
6	523.2	48.3	1062	16	AAT04050
7	523.2	48.3	1062	17	AAT34852
8	523.2	48.3	1062	17	AAT37383
9	523.2	48.3	1062	17	AAT32591
10	523.2	48.3	1062	18	AAT85555
11	523.2	48.3	1062	19	AAV21696

12	523.2	48.3	1062	21	AAAS1991	Human thrombopoietin
13	523.2	48.3	1086	17	AAAT47958	Human thrombopoietin
14	523.2	48.3	1342	16	AAQ99704	Human MGDF-1 and M
15	523.2	48.3	1342	17	AAAT36657	Native human mpl 1
16	523.2	48.3	1342	17	AAAT10025	Human MGDF-1/2 cDN
17	523.2	48.3	1721	16	AAQ99554	Thrombopoietin cod
18	523.2	48.3	1721	21	AAZ40191	Human wild type th
19	523.2	48.3	1774	21	AAAT47793	Human thrombopoiet
20	523.2	48.3	1795	16	AAQ94107	hML cDNA. Homo sa
21	523.2	48.3	1795	18	AAAT64318	Human thrombopoiet
22	521.6	48.2	600	17	AAAT36658	Truncated human mp
23	521.6	48.2	605	20	AAAX32813	Human truncated mp
24	521.6	48.2	1267	16	AAQ99553	Thrombopoietin cod
25	515.6	47.6	525	11	AAQ04482	Plasmid PASN6 enco
26	512.4	47.3	525	11	AAQ04484	Plasmid PASN145 en
27	507.6	46.9	525	11	AAQ04481	Plasmid PAS28 enco
28	500.2	46.2	1342	17	AAAT33933	Human megakaryocyt
29	500.2	46.2	1342	19	AAAT29068	Nucleotide sequenc
30	500.2	46.2	1342	20	AAAX32812	Human mpl ligand e
31	489.2	45.2	1050	17	AAAT41901	Human thrombopoiet
32	486.8	44.9	525	9	AAAN80947	G-CSF gene isolate
33	486.8	44.9	1520	8	AAAN70223	Plasmid pBRV2 inse
34	486.8	44.9	1520	8	AAAN71320	Sequence encoding
35	486.8	44.9	1521	7	AAAN60937	Plasmid pBRV2 inse
36	486.8	44.9	1525	9	AAAN81478	Sequence encoding
37	486.8	44.9	1525	10	AAAN91086	Plasmid pP12 contg
38	486.6	44.9	1017	16	AAQ97181	PMON15937 DNA enco
39	486.6	44.9	1017	21	AAAO3735	Human interleukin-
40	485.2	44.8	1415	8	AAAN71089	Sequence of human
41	479.8	44.3	741	17	AAAT41907	Human thrombopoiet
42	478	44.1	861	17	AAAT41902	Human thrombopoiet
43	478	44.1	861	17	AAAT15486	Megakaryocyte diff
44	478	44.1	946	17	AAAT29735	Human platelet pro
45	478	44.1	946	21	AAAZ44174	Human platelet-pro

ALIGNMENTS

RESULT 1
AAT41787
ID AAT41787 standard; DNA; 1083 BP.
XX
AC AAT41787;
XX
DT 01-JUL-1997 (first entry)
XX
DE Fusion peptide #2 having G-CSF and TPO activity.
XX
KW Fusion protein; human granulocyte colony stimulating factor; hG-CSF;
KW thrombopoietin; TPO; spacer peptide; blood platelet production;
KW leukocyte production; anaemia; ds.
XX
OS Homo sapiens.
XX
FH Key
FT sig_peptide Location/Qualifiers
FT mat_peptide 1..63
FT 64..1083 /*tag= a
FT /*tag= b
XX
PN WO9634016-A1.
XX
PD 31-OCT-1996.
XX
PF 26-APR-1996; 96WO-JP01157.
XX
PR 26-APR-1995; 95JP-0102625.
XX
PA (KYOW) KYOWA HAKKO KOGYO KK.
XX
PI Anazawa H, Konishi N, Shiotsu Y, Tamaoki T, Terasaki Y;
PI Uchida K, Yamasaki M, Yamashita K, Yokoi H;

XX WPI; 1996-497573/49.
 DR P-PSDB; AAW00378.
 XX
 XX Fusion peptide having G-CSF and with thrombopoietin activity -
 PT optionally chemically modified with a poly:alkylene glycol, used for
 PT treatment of anaemia
 XX
 XX Claim 3; Page 49-51; 71pp; Japanese.
 PS
 CC The sequences given in AAT41786-88 encode fusion proteins which are
 CC composed of a peptide having human granulocyte colony stimulating
 CC factor (hG-CSF) activity fused with a peptide having thrombopoietin
 CC (TPO) activity, opt. via a spacer peptide. Peptides derived from
 CC these by deletion, insertion or substitution of one or more amino
 CC acid residues are included within the scope of the invention. The
 CC fusion peptides stimulate blood platelet and leukocyte production
 CC and are useful in the treatment of anaemia.
 XX
 XX Sequence 1083 BP; 197 A; 355 C; 305 G; 226 T; 0 other;
 SQ

Query Match 100.0%; Score 1083; DB 17; Length 1083;
 Best Local Similarity 100.0%; Pred. No. 2.2e-249; Indels 0; Gaps 0;
 Matches 1083; Conservative 0; Mismatches 0;

QY 1 ATGGAGCTGACTGAATTGCTCTCGTGGTCATGCTTCTTAACCTCAAGGCTAAGCGTGC 60
 DB 1 atggagctgactgaattgctctcggttcacgtcttcttaactcaagctaaagctaaagctg 60

QY 61 TCACGCGCGCTCTCTGCTGTGTGACCTCCGAGTCCGAGTCTCAGTAAACTGCTTGTGACTGCC 120
 DB 61 tcacgcgcgctctctgctgtgtgacctccgagtcctcagtaaacctgtcttctgactcc 120

QY 121 CATGTCTCTCACAGCAGACTGAGCCAGTGCACGAGTTCACCCCTTTCGCTTACCTGCTC 180
 DB 121 catgtctctcacagcagactgagccagtgcacagaggttcaccccttctgcttaccctgtc 180

QY 181 CTGCTGCTGCTGTGGAGCTTTAGCTTGGGAGATGAAACCCAGATGGAGGAGACCAAG 240
 DB 181 ctgctgctgctgtggagcttttagcttgggagatgaaacccagatggagagagaccaaag 240

QY 241 GCACAGGACATTCGGAGCAGTACCCCTTCTGCTGGAGGAGTGTATGCGACACGCGGGA 300
 DB 241 gcacaggacatctggagcagtaccccttctgctggagagtgatggtgagcagcagggga 300

QY 301 CAATGGGACCCACTTGCCTCTCATCCCTCCCTGGGACGCTTTCGGACAGCTCCGCTC 360
 DB 301 caatgggacccacttgcctctcatccctccctgggagagcttctggacaggtccgctctc 360

QY 361 CTCCTTTGGGCCCCGACGAGCCTCTCTTGGAGCCAGCTTCTCCACAGGCGACGACCA 420
 DB 361 ctcccttggggccccgacgagcctctcttggagccagcttctccacagggcagagccaca 420

QY 421 GCTCACAGGATCCCAATGCCATCTTCCTGAGCTTCCATCAACCTGCTCCGAGGAAAGGTG 480
 DB 421 gctcacaggatcccaatgccatcttctgagcttccatcaacactgctccgaggaaggtg 480

QY 481 CGTTTCTGATGCTTAGAGGGTCCACCCCTTCCGCTCAGGGTGGCGTTCGAGGT 540
 DB 481 cgtttctgatgcttagaggggtccaccccttccgctcaggggtggcggttctgaggt 540

QY 541 GGTTCGGAGGGGTTCAGAGCACCACATATCGCGCTCCAGTCTACACAGAGCTTC 600
 DB 541 ggttcggaggggttcagagcaccacatatacgcgctccaggtctaccacagaggttc 600

QY 601 CTTTTAAAAGCTTAGAGCAAGTAGGAGATCCAGGGCGATGGCGCAGCGCTCCAGGAG 660
 DB 601 cttttaaaagcttagagcaagtaggagatccagggcgatggcgagcgctccagag 660

QY 661 AAGCTGTGTGCCACCTACAGCTGTGCCACCCGAGGAGCTGGTCTGCTGGACACTCT 720
 DB 661 aagctgtgtgccacctacagctgtgccacccgagagctgggtctgctggacactct 720

QY 721 CTGGGCATCCCTGGGCTCCCTGAGCAGCTGCCAGCCAGCAGGCTGCGAGCTGGCAGGC 780
 DB 721 ctgggcataccctgggctccctgagcagctgccagccagccagctgagctggcaggc 780

QY 781 TGCTTGAGCAACTCATAGCGGCTTTTCTTACACAGGGGCTCTGCGAGGCCCTGGAA 840
 DB 781 tgcttgagcaactcatagcgccctttctctaccaggggctctcgagggccctgga 840

QY 841 GGGATCTCCCGGAGTGGGTCCACCTTGGACACTGCGAGCTGGACGCTGCGGACTTT 900
 DB 841 gggatctcccgagtggttccaccttggacactggacactgagctggacgtcgccgacttt 900

QY 901 GCACACCACTCTGGCAGCAGATGGAAGAACTGGGAATGGCCCTCTGCCCTGCGAGCCAC 960
 DB 901 gcacaccactctggcagcagatggaagaactgggaatggccctgcccctgagccacc 960

QY 961 CAGGGTGCCATCGCGGCTTCGCTCTGCTTTCACAGCGCGGCGAGAGGGTCTTAGTT 1020
 DB 961 cagggtgccatcgcgcccttcgctctgcttctccagcgccgaggggtctctagtt 1020

QY 1021 GCCTCCCATCTGCAGAGCTTCTGGAGGTGCTGCTACCGGCTTCTAGCCACCTTGCCAG 1080
 DB 1021 gcctcccatctgcagagcttctggaggtgctgctaccgcttctacgcccacttgccag 1080

QY 1081 CCC 1083
 DB 1081 ccc 1083

RESULT 2
 AAT41788
 ID AAT41788 standard; DNA; 1095 BP.
 XX
 AC AAT41788;
 XX
 DT 01-JUL-1997 (first entry)
 XX
 DE Fusion peptide #3 having G-CSF and TPO activity.
 XX
 KW Fusion protein; human granulocyte colony stimulating factor; hG-CSF;
 KW thrombopoietin; TPO; spacer peptide; blood platelet production;
 KW leukocyte production; anaemia; ds.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT sig_peptide 1..63
 FT mat_peptide 64..1095
 FT mat_peptide /*tag= a
 FT mat_peptide /*tag= b
 XX
 PN WO9634016-A1.
 XX
 PD 31-OCT-1996.
 XX
 XX 26-APR-1996; 96WO-JP01157.
 PF
 XX 26-APR-1995; 95JP-0102625.
 PR
 XX (KYOW) KYOWA HAKKO KOGYO KK.
 PA
 XX Anazawa H, Konishi N, Shiotsu Y, Tamaoki T, Terasaki Y;
 PI Uchida K, Yamasaki M, Yamashita K, Yokoi H;
 XX
 XX WPI; 1996-497573/49.
 DR P-PSDB; AAW00379.
 XX
 XX Fusion peptide having G-CSF and with thrombopoietin activity -
 PT optionally chemically modified with a poly:alkylene glycol, used for
 PT treatment of anaemia
 XX
 PS Claim 3; Page 52-54; 71pp; Japanese.

XX The sequences given in AAT41786-88 encode fusion proteins which are
CC composed of a peptide having human granulocyte colony stimulating
CC factor (hG-CSF) activity fused with a peptide having thrombopoietin
CC (TPO) activity, opt. via a spacer peptide. Peptides derived from
CC these by deletion, insertion or substitution of one or more amino
CC acid residues are included within the scope of the invention. The
CC fusion peptides stimulate blood platelet and leukocyte production
CC and are useful in the treatment of anaemia.
XX
SQ Sequence 1095 BP; 194 A; 362 C; 311 G; 228 T; 0 other;

Query Match 95.1%; Score 1030.2; DB 17; Length 1095;
Best Local Similarity 97.7%; Pred. No. 9e-237;
Matches 1070; Conservative 0; Mismatches 13; Indels 12; Gaps 2;

QY 1 ATGAGCTGACTGAATGCTCTCGTGGTCATGCTTCTCTAATCAAGGCTTAAGCGTG 60
DB 1 atgagctgactgaattgctctcggtgctatgctctctcctaactgcaaggctaaagctg 60

QY 61 TCCAGCCCGGCTCTCTCTGCTTGTGACCTCCGAGTCTCAGTAAGTCTGCTGACTCC 120
DB 1 tccagcccggtctctctgcttgtgacctcccgagtcctcagtaaacgtctctgactoc 120

QY 121 CATGTCCTTACACAGCACTGACCCAGTGCACAGAGTTTACACCTTTGCCATACACCTGTC 180
DB 121 catgtctctcaagcaactgacagcagtgcccaagagttcacaccttgcctacacctgtc 180

QY 181 CTGCTGCCCTGCTGGACTTTAGCTTGGGAGATGGAACACCCAGATGGAGGACCAAG 240
DB 181 ctgctgccctgctggactttagcttggagatggaacacccagatggaggagaccaag 240

QY 241 GCACAGGACATTTGGGAGCAGTGCCTTCTGCTGAGGGAGTGATGCAGCACCGGGA 300
DB 241 gcacaggacattctgggagcagtgacctctctgctgaggagtgatgcaagcagcgga 300

QY 301 CAACTGGGACCCACTTGTCTCTCATCCCTCTGCTGGGAGCTTCTGGACAGGTCCTCTC 360
DB 301 caactgggaccacttgcctctcatccctctgctgggagcttcttgacaggtccgtctc 360

QY 361 CTCCTTGGGGCCCTGCAGAGCTTCCTTGAACCCAGCTTCTCCACAGGCGAGGACACA 420
DB 361 ctcttggggccctgcagagctctcttggaaacctctctccacagggcaggaacca 420

QY 421 GCTCACAGGATCCCAATGCCATCTTCTGAGCTTCCACACCTGCTCCGAGGAAAGGTG 480
DB 421 gctcacaggatcccaatgccatcttctgagcttcccaacacctgctccgaggaagggtg 480

QY 481 CGTTTCTGATGCTTAGGAGGGTCCACCTCTGCGTCAAGG---GGTGGCGGTTCTGGA 537
DB 481 cgttctctgatgcttagagaggtccacctctgctgacggtcccgaggtggcctcggc 540

QY 538 GGTGTTCCGAGGGGGTTC-----TAGACACCAACATATATCGCGCTCGAGTCTRA 598
DB 541 ggtggtctggtgagcgtccgagcggtcggtgcccacacatctcgcctcagtcata 600

QY 589 CCACAGAGCTTCTTTTAAAGCTTAGACAGTGCAGGAGATCCAGGGCGATGCGGCA 648
DB 601 ccacagagcttcttttaaagcttagagcaagtgaggaagatccagagcgatggcgca 660

QY 649 GCCTCCAGGAGAAGCTGTGTGCCACTACAAAGCTGTGCGCACCCCGAGGAGTGTGTG 708
DB 661 gcgctccaggaagctgtgtgccactacaagctgtgccacctccagagagctggtgctg 720

QY 709 CTCGGACACTCTTGGGATCCCTTGGGCTGCCCTGAGCAGTGTGCCAGCCAGGCGCTG 768
DB 721 ctggsacactctctgggactccctgggtccctgagcagctgcccagcagcagccctg 780

QY 769 CAGCTGGCAGGCTGTGAGCAACTTCATAGGCGCTTTTCTCTACACGGGCTCCG 828
DB 781 cagctggcaggtgtgtgagcaactccatagcgcccttctctaccaggggctcctg 840

QY 829 CAGGCCCTGGAAGGGATCTCCCCGAGTTGGTCCACCTTGGACACACTGCGAGCTGGAC 888
DB 841 cagggccctggaaggatctcccccgagttgggtccccaccttgacacactgcaagctggac 900

QY 889 GTCCGCGACTTTGCCACCACTCTGCGACGACATGGAAGAACTGGGAATGGCCCTGCC 948
DB 901 gtcgcgactttgccaccacctctgacagcagatggaagaactgggaatggccctgccc 960

QY 949 CTCAGCCCACTCCAGGAGTGCATGCGCGCTTCCCTCTCTCTTTCAGCGCCGGCAGGA 1008
DB 961 ctgagcccaactccaggggtgccatgcccgtccctgctctgtcttccagcgccggcagga 1020

QY 1009 GGGTCTCTAGTTCCTCCCATCTGCAGAGCTTCTCTGAGAGTGTGCTACCGCTTACGC 1068
DB 1021 ggggtctctagttgctcccatctgcagagcttctctgaggtgctgacgcggtctcagc 1080

QY 1069 CACCTTGGCCAGGCC 1083
DB 1081 caccttggccagccc 1095

RESULT 3
AAT41786
ID AAT41786 standard; DNA; 1047 BP.
XX
AC AAT41786;
XX
DT 01-JUL-1997 (first entry)
XX
DE Fusion peptide #1 having G-CSF and TPO activity.
XX
KW Fusion protein; human granulocyte colony stimulating factor; hG-CSF;
KW thrombopoietin; TPO; spacer peptide; blood platelet production;
KW leukocyte production; anaemia; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT sig_peptide 1..63
FT /*tag= a
FT mat_peptide 64..1047
FT /*tag= b
XX
PN W09634016-A1.
XX
PD 31-OCT-1996.
XX
PF 26-APR-1996; 96WO-JP01157.
XX
PR 26-APR-1995; 95JP-0102625.
XX
PA (KYOW) KYOWA HAKKO KOGYO KK.
PI Anazawa H.; Konishi N, Shiotsu Y, Tamaoki T, Terasaki Y;
PI Uchida K, Yamasaki M, Yamashita K, Yokoi H;
XX
DR WPI: 1996-497573/49.
XX P-PSDB; AAW00377.
PT Fusion peptide having G-CSF and with thrombopoietin activity -
PT optionally chemically modified with a poly:alkylene glycol, used for
PT treatment of anaemia
XX
PS Claim 3; Page 46-48; 71pp; Japanese.
XX
CC The sequences given in AAT41786-88 encode fusion proteins which are
CC composed of a peptide having human granulocyte colony stimulating
CC factor (hG-CSF) activity fused with a peptide having thrombopoietin
CC (TPO) activity, opt. via a spacer peptide. Peptides derived from
CC these by deletion, insertion or substitution of one or more amino
CC acid residues are included within the scope of the invention. The
CC fusion peptides stimulate blood platelet and leukocyte production
CC and are useful in the treatment of anaemia.
CC

```

XX
SQ Sequence 1047 BP; 192 A; 351 C; 288 G; 216 T; 0 other;

Query Match          91.7%; Score 993; DB 17; Length 1047;
Best Local Similarity 96.2%; Pred. No. 6.8e-228;
Matches 1042; Conservative 0; Mismatches 5; Indels 36; Gaps 1;

QY 1 ATGGAGCTGACTGAATGCTCCCTGGTGCATGCTTCTCTAACTCAAGGCTTAAGCGTGG 60
DB 1 atggagctgactgaattgctccctcggtgcattgctctcttaactcaaggctcaacgctg 60

QY 61 TCCAGCCCGGCTCCTCTGCTTGTGACCTCCGAGTCTCAGTAAACTGCTTGTGACTCC 120
DB 61 tccagcccggtcctctgcttgtgacctccgagtcctcagtaaaactgcttgactcc 120

QY 121 CATGCTCTCACAGCAGACTGAGCCAGTCCCGAGAGGTTCAACCTTTGCTACACCTGTC 180
DB 121 catgctctcacagcagactgagccagtcccgagaggttcacaccttgcctacacctgc 180

QY 181 CTGCTGCTGCTGTGGACTTTAGCTTGGGAGATGAAACCCAGATGGAGGAGCAAG 240
DB 181 ctgctgctgctgtggacttttagcttgggagatgaaacccagatggagagaccaag 240

QY 241 GCACAGACATCTGGGAGCAGTGACCTTCTGCTGGAGGAGTGATGACAGCAGGGGA 300
DB 241 gcacagacatctgggagcagtgaccttctgctggagggagtgaatgscagcagggga 300

QY 301 CAACTGGGACCACTTGCCTCTCATCCCTGCTGGGCGACGTTTCTGGACAGGTCCGTC 360
DB 301 caactgggacccacttgcctctcatccctgctgggagcttctggacaggtccgctc 360

QY 361 CTCTCTGGGCGCTGAGAGCCCTTCTGGAACCCAGCTTCTCCACAGGCGAGGCCACA 420
DB 361 ctctctgggcgctgagagcccttctggaacccagcttctccacagggcgagccaca 420

QY 421 GCTCACAGGATCCCAATGCTTCTTCTGAGCTTCCAACTGCTCCGAGGAAAGTGTG 480
DB 421 gctcacaggatcccaatgcttcttctgagcttccaaactgctccgaggaagtgtg 480

QY 481 CGTTTCTGATGTTGTAGGAGGTCACCCCTCTGCTGAGGGGTGGCGGTTCTGGAGGT 540
DB 481 cgtttctgatgttgttaggaggtcacccctctgctgaggggtggcggttctggaggt 540

QY 541 GGTTCGGAGGGGTCTTAGAGCAACCATATCGGGCTCGAGTCTACACAGAGCTTTC 600
DB 541 ggttcggaggggtcttagagcaacccatctcgggctcgagtctacacagagcttctc 600

QY 601 CTTTTAAAAAGCTTAGAGCAAGTAGAGATCCAGGGCGATGGCGAGCGCTCCAGGAG 660
DB 601 cttttaaaaagcttagagcaagtagagatccagggcgatggcgagcgctccaggag 660

QY 624 CTTTAAAGCTTAGAGCAAGTAGAGATCCAGGGCGATGGCGAGCGCTCCAGGAG 624
DB 624 cttttaaagcttagagcaagtagagatccagggcgatggcgagcgctccaggag 624

QY 661 AAGCTGTGTCACCTTACAGCTGTCCACCCGAGGAGCTGGTGTGCTCGGACACTCT 720
DB 661 aagctgtgtgacacttacagctgtgccacccgagagctggtgtgctcggaactct 720

QY 625 AAGCTGTGTCACCTTACAGCTGTCCACCCGAGGAGCTGGTGTGCTCGGACACTCT 684
DB 625 aagctgtgtgacacttacagctgtgccacccgagagctggtgtgctcggaactct 684

QY 721 CTGGGATCCCTGGGCTCCCTGAGCAGCTGCCCGAGGCGCTGAGCTGGCAGGCG 780
DB 721 ctgggatccctgggctccctgagcagctgcccgagggcgctgagctggcaggcg 780

QY 744 CTGGGATCCCTGGGCTCCCTGAGCAGCTGCCCGAGGCGCTGAGCTGGCAGGCG 744
DB 744 ctgggatccctgggctccctgagcagctgcccgagggcgctgagctggcaggcg 744

QY 781 TGCTTGAGCCAACTCCATAGCGGCTTTTCTCTATACAGGGGCTCTGAGGCGCTGGAA 840
DB 781 tgcttgagccaaactccatagcggttttctctatacaggggctctgagggcgctggaa 840

QY 841 GGGATCTCCCGGAGTTGGTCCACCTTGGACACTGACACTGGAGCTGCCCGACATTT 900
DB 841 gggatctcccgagttgggtccaccttggacactgacactggagctgcccgacattt 900

QY 805 GGGATCTCCCGGAGTTGGTCCACCTTGGACACTGACACTGGAGCTGCCCGACATTT 864
DB 805 gggatctcccgagttgggtccaccttggacactgacactggagctgcccgacattt 864

QY 901 GCCACCACTATCGCAGCAGATGGAAGAACTGGGAATGGCCCTTCCCTGACGCCACCC 960
DB 901 gccaccactatcgccagcagatggaagaaactgggaatggcccttccctgacgccacc 960

QY 865 GCCACCACTATCGCAGCAGATGGAAGAACTGGGAATGGCCCTTCCCTGACGCCACCC 924
DB 865 gccaccactatcgccagcagatggaagaaactgggaatggcccttccctgacgccacc 924

QY 961 CAGGGTGCCATGCCGCGCTTCCGCTCTGCTGTTTCCAGCGCCGGGAGGAGGCTCTAGTT 1020
```

```

DB 925 cagggtgccatgccggccttcgctctgtttccagcgccggcagggggtcctagt 984
QY 1021 GCCTCCCATCTGCAGAGCTTCTGGAGGTGTCGTACCGGCTTCTACGCCACCTTGCCAG 1080
DB 985 gctcccatctcagagcttctcctggaggtgtgtaccggttctacgcccaccttgcccag 1044
QY 1081 CCC 1083
DB 1045 ccc 1047

RESULT 4
AAQ99552
ID AAQ99552 standard; cDNA to mRNA; 861 BP.
XX AAQ99552;
XX 29-APR-1996 (first entry)
XX Thrombopoietin coding sequence.
DE Thrombopoietin; TPO; humoral factor; platelet; antibody; therapy; AIDS;
KW thrombocytopenia; hypoplastic anaemia; thrombotic thrombocytopenia;
KW disseminated intravascular coagulation syndrome; ds.
XX Homo sapiens.
XX Key Location/Qualifiers
FH CDS 25..783
FT /*tag= a
FT /*product= thrombopoietin
FT sig_peptide 25..87
FT /*tag= b
FT mat_peptide 88..783
FT /*tag= c
XX WO9521919-A2.
XX 17-AUG-1995.
XX 14-FEB-1995; 95WO-JP00208.
XX 14-FEB-1994; 94JP-0039090.
XX 14-MAR-1994; 94US-0212164.
XX 25-MAR-1994; 94JP-0079842.
XX 01-APR-1994; 94US-0221020.
XX 01-JUN-1994; 94JP-0155126.
XX 15-JUN-1994; 94JP-0167328.
XX 20-JUL-1994; 94US-0278083.
XX 17-AUG-1994; 94JP-0193169.
XX 18-AUG-1994; 94JP-0193916.
XX 11-OCT-1994; 94US-0320300.
XX 01-NOV-1994; 94JP-0304167.
XX 01-DEC-1994; 94JP-0298669.
XX 22-DEC-1994; 94US-0361811.
XX 28-DEC-1994; 94JP-0341200.
XX 31-JAN-1995; 95US-0381478.
XX 17-AUG-1994; 94JP-0227159.
XX (KIRI ) KIRIN BREWERY KK.
XX Akahori H. Iwamatsu A. Kato T, Kuroki R, Miyazaki H;
PI Muto T, Ohgami K, Shimizu T;
XX WPI; 1995-293120/38.
DR P-PSDB; AAR81378.
XX Thrombopoietin polypeptide which specifically stimulates or
PT increases platelet production - useful in treatment of
PT thrombocytopenia, also related DNA and vectors
XX Claim 19; Page 337-338; 383pp; English.
```


Qy	361	CTCCTTGGGGCCCTGCAGAGCCTCTCTTGAACCCAGCTTCTCCACAGGCGAGACACA	420
Db	385	ctccttggggccctgcagagctccttggaaaccagcttctccacagggcaggaccaca	444
Qy	421	GCTCACAAGATCCCAATGCCATCTCTCTGAGCTTCCAAACCTGCTCCGAGGNAAGGTG	480
Db	445	gctccaaaggatcccaatgccattctcttgagcttcccaacacctgcctccaggaagaggtg	504
Qy	481	CGTTTCTCATGCTGTAGGAGGGTCCACCCCTCTCGCTCAGGGGTGGC	528
Db	505	cgttctctgatctgttagagaggtccaccctctgcgtcaggcgggcc	552
RESULT	5		
AAT03941			
ID	AAT03941	standard; cDNA; 1062 BP.	
XX			
AC	AAT03941;		
XX			
DT	28-APR-1996	(first entry)	
XX			
DE	Human thrombopoietin coding sequence.		
XX			
KW	Thrombopoietin; erythropoiesis stimulator; treatment;		
KW	thrombocytopenia; anaemia; ds.		
XX			
OS	Homo sapiens.		

```

QY 361 CTCCTTGGGGCCCGCAGAGCCCTCTTGGAAACCCAGCTTCTCCACAGGCGAGACCACA 420
Db 361 ctccctggggccctcgagagcctctcttggaaccagcttctccacagggcagaccaca 420
QY 421 GCTCACAGAGTCCCAATGCATCTCTCTGAGCTTCCACACACTGCTCCGAGGAAAGGTG 480
Db 421 gctcacagatcccaatgcacatctctctgagcttccacacactgctccgaggaaggtg 480
QY 481 CGTTTCTGATGCTGTAGGAGGTCACACCTCTGCTGCTGAGGGGTGC 528
Db 481 cgtttctgatgctgtcagaggggtccaccctctgctcagcgggcc 528

RESULT 6
AAT04050
ID AAT04050 standard; cDNA; 1062 BP.
XX
AC AAT04050;
DT 28-MAR-1996 (first entry)
DE Sequence encoding haematopoietic protein.
XX
KW Haematopoiesis; protein; bone marrow; stem cells; precursor;
KW platelet; therapy; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS l..1062
FT FT /*tag= a
FT FT /product= Haematopoietic protein.
XX
PN WO9521920-A1.
XX
PD 17-AUG-1995.
XX
XX 05-AUG-1994; 94WO-US08806.
XX
PR 01-JUN-1994; 94US-0252491.
PR 14-FEB-1994; 94US-0196025.
PR 25-FEB-1994; 94US-0203197.
PR 21-MAR-1994; 94US-0215203.
XX
XX (UNIW ) UNIV WASHINGTON.
XX (ZYMO ) ZYMOGENETICS INC.
XX
PI Burkhead SK, Foster DC, Hagen FS, Holly RD, Kaushansky K;
PI Kuijper JL, Lofton-day C, Lok S, Oort PJ;
XX
XX WPI; 1995-293121/38.
XX P-PSDB; AAR79906.
XX
XX Haematopoietic proteins and polypeptide(s) - useful for in vivo and
XX ex vivo therapy
XX
XX Claim 20; Page 106-108; 137pp; English.
XX
XX Haematopoietic proteins and polypeptides are useful for stimulating
XX platelet production in a mammal when given in a therapeutically
XX effective amount. They are also useful for stimulating bone marrow
XX cell proliferation where the bone marrow cells are megakaryocytes or
XX their precursors.
XX
XX Sequence 1062 BP; 223 A; 375 C; 236 G; 228 T; 0 other;

Query Match 48.3%; Score 523.2; DB 16; Length 1062;
Best Local Similarity 99.4%; Pred. No. 9.8e-116;
Matches 525; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGGAGCTGACTGAATGTCTCTCTGCTGGTCATGCTTCTCTTAACGTCAAGGCTAAGCGTG 60

```

```

Db 1 atggagctgactgaattgctcctcctcggtcatgctctctcttaactgcaaggctaaacgtg 60
QY 61 TCACAGCCCGGCTCCTCCTGCTTGTGACCTCCGAGTCTCTAGTAACCTGCTTCTGACTCC 120
Db 61 tcacagcccggtcctcctcctcctgctgtgacctccgagtcctcagtaaacctgctcgtgactcc 120
QY 121 CATGTCTTTCACAGCAGACTGAGCCAGTGCACAGAGGTTTCCACCTTTTGCCTACACCTGTC 180
Db 121 catgtcttccacagcagactgagcagtgcccagaggttccacctttgctacacactgtc 180
QY 181 CTGCTGCCCTGCTGTGACTTTAGCTTGGGAGAAATGGAACCCAGATGAGGAGACCAAG 240
Db 181 ctgctgctgctgctggaacttagcttggaatggaaatgaaacccagatggagggagaccaag 240
QY 241 GCACAGGACATTTCTGGGAGCAGTGACCCCTCTCTGCTGGAGGAGTGATGGCAGCACGGGA 300
Db 241 gcacaggacattctggagcagtgaccttctgctggagggagtgatggcagcacggga 300
QY 301 CAACTGGGACCCACTTGCCTCTCATCCCTCTGGGCGAGCTTTCTGACAGGTCCTGCTC 360
Db 301 caactgggacccacttgcctctcatccctcctcctcctgggagcagtttctggacaggtccgtctc 360
QY 361 CTCTTGGGGCCCTGCAGAGCCTCTTGGAAACCCAGCTTCTCTCCACAGGCGAGACCACA 420
Db 361 ctcttggggccctgcagagcctccttggaaacccagcttctccacacaggttccacagggcagaccaca 420
QY 421 GCTCACAGGATCCCAATGCCATCTTCTGAGCTTCCACACCTGCTCCGAGGAAAGGTG 480
Db 421 gctcacaggatcccaatgccatcttctcagcttccacacactgctccgaggaaggtg 480
QY 481 CGTTTCTGATGCTGTAGGAGGTCACACCTCTGCTGCTGAGGGGTGC 528
Db 481 cgtttctgatgctgtgtagaggggtccacctctgctcagcgggcc 528

RESULT 7
AAT34852
ID AAT34852 standard; cDNA; 1062 BP.
XX
AC AAT34852;
XX
DT 03-DEC-1996 (first entry)
XX
XX Human thrombopoietin cDNA.
XX
XX Thrombopoietin; TPO; thrombocytopenia; platelet; therapy; ds.
XX
XX Homo sapiens.
XX
XX WO9617062-A1.
XX
XX 06-JUN-1996.
XX
XX 15-NOV-1995; 95WO-US14929.
XX
XX 30-NOV-1994; 94US-0346999.
XX
XX (ZYMO ) ZYMOGENETICS INC.
XX
XX Buddle MM, Downey W, Forstrom JW, Foster DC, Lok S;
XX Osborn SG;
XX
XX WPI; 1996-286830/29.
XX P-PSDB; AAR98947.
XX
XX New low mol. wt. forms of thrombopoietin and related DNA - vectors
XX and transformed cells, useful for stimulating platelet prodn. esp.
XX for treatment of thrombocytopenia
XX
XX Claim 19; Page 57-59; 84pp; English.
XX
XX A cDNA clone (AAT34852) codes for human thrombopoietin (AAR98947), a

```



```
RESULT 9
AAT32591
ID AAT32591 standard; cDNA; 1062 BP.
XX
XX AC AAT32591;
XX
XX 30-SEP-1996 (first entry)
XX
XX DE Human thrombopoietin cDNA.
XX
XX KW Thrombopoietin; TPO; tissue plasminogen activator; cytokine;
XX KW protein secretion; signal peptide; thrombocytopenia; therapy; ds.
XX
XX OS Homo sapiens.
XX
XX FH Key Location/Qualifiers
XX sig_peptide 1..63
XX FT /*tag= a
XX FT 64..1059
XX FT mat_peptide /*tag= b
XX
XX PN WO9617067-A1.
XX
XX PD 06-JUN-1996.
XX
XX PF 15-NOV-1995; 95WO-US14932.
XX
XX PR 30-NOV-1994; 94US-0347029.
XX
XX PA (ZYMO ) ZYMOGENETICS INC.
XX
XX PI Foster DC, Heipel MD, Holly RD;
XX
XX DR WPI; 1996-277784/28.
XX
XX DR P-PSDB; AAR99599.
XX
XX PT DNA construct encoding native mammalian tissue plasminogen activator
XX PT secretory peptide and thrombopoietin polypeptide - used to produce
XX PT thrombopoietin polypeptide(s) on a large scale and in cost effective
XX PT manner
XX
XX PS Disclosure; Page 41-43; 56pp; English.
XX
XX CC A cDNA clone (AAT32591) codes for human thrombopoietin (TPO) (AAR99599).
XX CC Novel DNA constructs code for a secretory peptide (AAR9600) fused to
XX CC the TPO polypeptide, with a cleavage site at the junction. Such
XX CC constructs can be expressed in transformed host cells, pref.
XX CC Pichia pastoris, Saccharomyces cerevisiae or rodent kidney cells,
XX CC for large-scale, cost-effective prodn. of human TPO, useful for the
XX CC treatment of thrombocytopenia.
XX
XX SQ Sequence 1062 BP; 223 A; 375 C; 236 G; 228 T; 0 other;

Query Match 48.3%; Score 523.2; DB 17; Length 1062;
Best Local Similarity 99.4%; Pred. No. 9.8e-116;
Matches 525; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ATGAGCTGACTGAATGCTCTGCTGTCATGCTTCTCTTAAGTCAAGGCTAAGCTG 60
Db 1 atggagctgactgaattgctctcgtggtcatgcttctcttaactgcaaggctaacgctg 60
Qy 61 TCAGCCCGGCTCTCTCTGCTGTCGACCTCCGAGTCTCTCAGTAAACTGCTTCGTGACTCC 120
Db 61 tcagcccggtctctctctgctgctgacctccagtcctcagtaaaactgcttctgactcc 120
Qy 121 CATGCTCTTCACAGCAGACTGACCCAGTGCACAGAGGTTTCACCCCTTTCCTACACCTGTC 180
Db 121 catgctcttcacagcagactgagccagtgccagaggttcaccccttgcctacacctgtc 180
Qy 181 CTGCTGCTGCTGCTGACTTTAGCTTGGGAGATGGAAACCCAGATGGAGGAGACCAG 240

Db 181 ctgctgctgctgctgacttttagcttggagaaatggaaaccagatggagagaccaag 240
Qy 241 GCACAGGACATTTCTGGAGCAGTACCCCTTCTGCTGGAGGGAGTGTATGTCAGACGGGGA 300
Db 241 gcacaggacattctggagcagtgacccttctgctggagggagtgatggcagcacggga 300
Qy 301 CAACCTGGGACCCACTTGCTCTCATCCCTCTGCTGGGACGCTTCTCTGACAGGTCCTGTC 360
Db 301 caactgggacccactgctctcatccctctcctcctgggagcagcttcttctgagcaggtccgtctc 360
Qy 361 CTCTCTTGGGGCCCTGCAGAGCCTCTTGAACCCAGCTTCTCCACAGGCGCAGGACCACA 420
Db 361 ctctctggggccctgcagagcctctcttgaaacccagctctctccacagggcagaccaca 420
Qy 421 GCTCACAGGATCCCAATGCCATCTTCTGAGCTTCCACACACTGCTCCGAGGAAAGGTG 480
Db 421 gctcaagagatcccaatgccatctctctgagcttcccaacactgctccgaggaagggtg 480
Qy 481 CGTTTCTGATGCTTGTAGGAGGGTCCACCCCTCTGCGTCAGGGGTGGC 528
Db 481 cgttctctgctgcttctgtaggaggtccaccctctcgtcagcggggccc 528

RESULT 10
AAT85555
ID AAT85555 standard; cDNA; 1062 BP.
XX
XX AC AAT85555;
XX
XX DT 05-NOV-1997 (first entry)
XX
XX DE Human thrombopoietin cDNA.
XX
XX KW TPO; cancer; myeloid cell proliferation; expansion; bone marrow;
XX KW peripheral blood stem cell; chemotherapy; radiation therapy;
XX KW breast cancer; leukaemia; lymphoma; multiple myeloma; ds.
XX
XX OS Homo sapiens.
XX
XX FH Key Location/Qualifiers
XX CDS 1..1062
XX FT /*tag= a
XX FT 64..1059
XX FT mat_peptide /*tag= b
XX FT /product= Thrombopoietin
XX
XX PN WO9640218-A1.
XX
XX PD 19-DEC-1996.
XX
XX PF 29-MAY-1996; 96WO-US07880.
XX
XX PR 07-JUN-1995; 95US-0482212.
XX
XX PA (ZYMO ) ZYMOGENETICS INC.
XX
XX PI Fibbe WE, Grossmann A;
XX
XX DR WPI; 1997-108626/10.
XX DR P-PSDB; AAW26841.
XX
XX PT Increasing haematopoietic cells in patient - by admin. of stem cells
XX PT from donor previously treated with thrombopoietin, used in patients
XX PT who have received chemotherapy or radiation for cancer
XX
XX PS Disclosure; Page 17-19; 32pp; English.
XX
XX CC In a claimed method, thrombopoietin (TPO) (especially human TPO) is
XX CC administered to a donor to stimulate proliferation of myeloid cells.
XX CC Bone marrow or peripheral blood stem cells are collected from the
XX CC donor and administered to the recipient. The method is used for
XX CC preparing cells for (bone marrow) transplantation and for stimulating
```

CC platelet or erythrocyte recovery in a patient receiving chemotherapy
CC or radiation therapy for e.g. breast cancer, leukaemia, lymphoma or
CC multiple myeloma. Treating the donor with TPO accelerates restoration
CC of erythrocyte and thrombocyte levels in the patient after transplant.
CC This restoration may be further improved by administering TPO to the
CC recipient after transplant. The present sequence encodes human TPO.
XX
SQ Sequence 1062 BP; 223 A; 375 C; 236 G; 228 T; 0 other;

Query Match 48.3%; Score 523.2; DB 18; Length 1062;
Best Local Similarity 99.4%; Pred. No. 9.8e-116;
Matches 525; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ATGGAGCTGACTGAATTCCTCCTCGTGTGCTATGCTTCTTAAGTGAAGCTAACGCTG 60
Db 1 atggagctgactgaattgctcctcggtgctatgcttcttaagtgaaagctaacgctg 60
Qy 61 TCCAGCCCGGCTCCTCGTGTGCTGACCTCCGAGTCCCTCAGTAACTGCTTCGTGACTCC 120
Db 61 tccagcccggtcctcctcggtgctgaccccgagtcctcagtaaaactgcttcgactcc 120
Qy 121 CATGCTTTCACAGACTGAGGAGTGCAGGAGTGCAGGAGTTCACCTTTCCTACACTGTC 180
Db 121 catgcttctcacagactgagcagtgccagaggttcaccccttgcctacactgtc 180
Qy 181 CTGCTGCTGCTGTGGACTTTAGCTTGGAGATGGAATGGAATGGAATGGAATGGAATGGA 240
Db 181 ctgctgctgctgtggacttttagcttggagatggaaatggaaatggaaatggaaatggaa 240
Qy 241 GCACAGGACATTCCTGGAGCAGTGAACCTTCTGCTGGAGGAGTGAATGGAATGGAATGGA 300
Db 241 gcacaggacattctggagcagtgacccttctgctggagagtgatggagcagcgggga 300
Qy 301 CAATGGGACCACTGCTCCTATCCCTCCTGGGAGCTTCCTGACAGGCTTCGACAGCTC 360
Db 301 caactgggaccacttgctctcctcctcctcctcctcctcctcctcctcctcctcctcct 360
Qy 361 CTCTTGGGGCCCTGACAGGCTCCTTGGAACTTCTTCCACAGGCTTCTTCCACAGGCTC 420
Db 361 ctcttggggccctgacagcctccttggaaacttcttccacagcttctcctcctcctcctc 420
Qy 421 GCTCACAAGGATCCCAATGCCATCTTCTGAGCTTCCAAACACTTCTCCGAGGAAAGTG 480
Db 421 gctcacaaggatcccaatgccatcttctgagcttccaaaccttctccagaggaagtg 480
Qy 481 CGTTTCTGATGCTGTGAGGAGGTTCACCTCCTGCGTTCAGGCTGCG 528
Db 481 cgtttctgatgctgtgaggggttccacctctcgtcagggcgggcc 528

RESULT 11
AAV21696
ID AAV21696 standard; cdna; 1062 BP.

AC AAV21696;
XX
XX
DT 07-AUG-1998 (first entry)
XX Human thrombopoietin encoding cdna.
DE
XX Thrombopoietin; TPO; expression vector; platelet number; cytokine;
KW human; bone marrow; proliferation; treatment; cancer; aplastic anaemia;
KW myelodisplastic syndrome; chemotherapy; cytopenia; thrombocytopenia;
KW haematologic disorder; leukaemia; lymphoma; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..1062
FT FT /*tag= a
FT FT /product= "Thrombopoietin polypeptide"
FT FT 1..63

FT mat_peptide /*tag= b
FT 64..1059 /*tag= c
XX
PN W09806849-A1.
XX
PD 19-FEB-1998.
XX
XX 30-JUL-1997; 97WO-US13543.
XX
PR 13-AUG-1996; 96US-0696447.
XX
PA (ZYMO) ZYMOGENETICS INC.
XX
PI Irani M, Morrison-nelson GR;
XX
XX WPI; 1998-159541/14.
DR P-PSDB; AAW53124.
XX
XX Thrombopoietin protein expression vector - used for increasing
PT platelet number in a mammal
XX
PS Disclosure; Pages 29-32; 56pp; English.
XX
CC This cdna encodes a thrombopoietin (TPO) polypeptide. This can be used
CC in the construction of a new expression vector replicable in a
CC eukaryotic host cell encoding TPO polypeptides. The vector comprises a
CC transcription promoter, a first DNA segment encoding a secretory leader,
CC a second segment encoding a TPO polypeptide and a transcription
CC terminator which are all operably linked. The second TPO segment consists
CC of C-X-B, where C is a human TPO cytokine domain, X is a peptide bond or
CC a linker consisting of one or two amino acid residues, such that X along
CC in combination with C or B does not provide a dibasic amino acid pair
CC and B is a polypeptide that can be selected from AAW53125 to AAW53126.
CC The secretory leader is a S. cerevisiae alpha-factor secretory leader. A
CC cultured yeast cell containing such an expression vector can be used to
CC produce the TPO polypeptide. The TPO polypeptide can be used in a method
CC for increasing platelet number in a mammal. It can be used to increase
CC proliferation of bone marrow cells for treatment of cytopenia, including
CC those induced by aplastic anaemia, myelodisplastic syndromes,
CC chemotherapy or congenital cytopenias. It can also be used to treat
CC thrombocytopenia, haematologic disorders, such as leukaemia and lymphoma
CC or metastatic cancers involving bone marrow.
XX
SQ Sequence 1062 BP; 223 A; 375 C; 236 G; 228 T; 0 other;

Query Match 48.3%; Score 523.2; DB 19; Length 1062;
Best Local Similarity 99.4%; Pred. No. 9.8e-116;
Matches 525; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 ATGGAGCTGACTGAATTCCTCCTCGTGTGCTATGCTTCTTCTTAAGTGAAGCTAACGCTG 60
Db 1 atggagctgactgaattgctcctcggtgctatgcttcttaagtgaaagctaacgctg 60
Qy 61 TCCAGCCCGGCTCCTCGTGTGCTGACCTCCGAGTCCCTCAGTAACTGCTTCGTGACTCC 120
Db 61 tccagcccggtcctcctcggtgctgaccccgagtcctcagtaaaactgcttcgactcc 120
Qy 121 CATGCTTTCACAGACTGAGCAGTGCAGGAGTTCACCTTTCCTACACTGTC 180
Db 121 catgcttctcacagactgagcagtgccagaggttcaccccttgcctacactgtc 180
Qy 181 CTGCTGCTGCTGTGACTTTAGCTTGGAGATGGAATGGAATGGAATGGAATGGAATGGA 240
Db 181 ctgctgctgctgtggacttttagcttggagatggaaatggaaatggaaatggaaatggaa 240
Qy 241 GCACAGGACATTCCTGGAGCAGTGAACCTTCTGCTGGAGGAGTGAATGGAATGGAATGGA 300
Db 241 gcacaggacattctggagcagtgacccttctgctggagagtgatggagcagcgggga 300
Qy 301 CAATGGGACCACTGCTCCTATCCCTCCTGGGAGCTTCCTGACAGGCTTCGACAGCTC 360
Db 301 caactgggaccacttgctctcctcctcctcctcctcctcctcctcctcctcctcctcct 360

Db 301 caactggaccacattgctctatccctctctggtggcagctttttgtgacagggtccctctc 360
 QY 361 CTCCTTGGGGCCCTGCAGAGCTCTCTTGAACCCAGCTTCTCCTCACAGGCGAGGACACCA 420
 Db 361 ctecttggggccctgcagagcctcttggaaaccagcttctctccacagggcagaccaca 420
 QY 421 GCTCACAGGATCCCATGCGATCTCTCTGAGCTTCCACACCTGCTCCGAGGAAGGTG 480
 Db 421 gctcacaggatcccaatgcatctctgagcttccacacacgtctccgaggaagggtg 480
 QY 481 CGTTTCTCTGATGCTGTGAGGAGGTCCACCTCTGCGTCAAGGGGTGGC 528
 Db 481 cgtttctctgatgctgttaggagggccaccctctgctcgagggcggcc 528

RESULT 12
 AAA51991
 ID AAA51991 standard; cDNA; 1062 BP.
 XX
 AC AAA51991;
 DT 04-DEC-2000 (first entry)
 XX
 DE Human thrombopoietin cDNA.
 XX
 KW Thrombopoietin; TPO; erythropoietin; EPO; erythropoiesis; stimulate;
 KW anaemia; radiation; chemotherapy; red blood cell production;
 KW haematopoietic cell; cytopenia; ex vivo expansion; anti-anaemic;
 KW immunomodulator; ds.
 XX
 OS Homo sapiens.
 XX
 Key Location/Qualifiers
 FT 1..1062
 FT CDS /*tag= a
 FT /product= Thrombopoietin
 XX
 PN US099830-A.
 XX
 PD 08-AUG-2000.
 XX
 PF 12-SEP-1997; 97US-0999927.
 XX
 PR 09-AUG-1994; 94US-0288417.
 PR 07-NOV-1994; 94US-0335366.
 XX
 PA (UNIW) UNIV WASHINGTON.
 PA (ZYMO) ZYMOGENETICS INC.
 XX
 PI Kaushansky K;
 XX
 WPI; 2000-557618/51.
 DR P-PSDB; AAY97136.
 XX

Stimulating erythropoiesis e.g. for treating anemias or cytopenia, or for ex vivo expansion of marrow or peripheral blood progenitor cell, comprises administering thrombopoietin alone or in combination with erythropoietin

Disclosure; Column 19-22; 23pp; English.

XX
 CC Erthropoiesis may be stimulated in anaemic mammals by administration of thrombopoietin (TPO), and optionally erythropoietin (EPO). The anaemic mammal to be treated will have shown a haematocrit level less than 33 percent of total blood volume, a reticulocyte count of less than 0.8 percent and may have been treated with radiation or chemotherapy. The improved red blood cell level is due to an increase in red blood cell production rather than a decrease in red blood cell destruction. TPO proteins may be used therapeutically to increase proliferation of haematopoietic cells in the bone marrow, such as in the treatment of cytopenia or anaemia, especially associated with bone marrow failure, disorders associated with low red blood cell production, particularly when accompanied by thrombocytopenia and for increasing the level of

CC circulating erythrocytes and erythrocyte precursor cells. TPO and EPO may be used for ex vivo expansion of marrow or peripheral blood progenitor cells.
 XX
 SQ Sequence 1062 BP; 223 A; 375 C; 236 G; 228 T; 0 other;

Query Match 48.3%; Score 523.2; DB 21; Length 1062;
 Best Local Similarity 99.4%; Pred. No. 9.8e-116;
 Matches 525; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGGAGCTGACTGAATTGCTCCTCGTGGTCATGCTTCTTAAGTCAAGGCTTAACGCTG 60
 Db 1 atggagctgactgaattgctcctcggtgctctctctctctctctctcaagctaaacgctg 60
 QY 61 TCCAGCCCGGCTCCTCTGCTTGTGACCTCCGAGTCTCTAGTAAACGCTTCGTGACTCC 120
 Db 61 tccagcccggtcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 120
 QY 121 CATGTCTTTCACAGCAGACTGACCCAGTCCCGAGAGGTTCACCCCTTTGCTACACCTGTC 180
 Db 121 catgtctttcacagcagactgagccagtgccacaggttcaccccttgcctacacctgtc 180
 QY 181 CTGCTGCTGCTGTGACTTTAGCTTTGGGAGATGAAACCCAGATGGAGGAGACCAAG 240
 Db 181 ctgctgctgctgtgacttttagcttggagaatgaaacccagatggaggagaccaag 240
 QY 241 GCACAGGACATTCCTGGAGCAGTGACCTTCTCTGTGAGGAGTGTATGGCAGCACGGGA 300
 Db 241 gcacaggacattcctggagcagtgacccttctcctgagggagtgatgagcagcagggga 300
 QY 301 CAACCTGGAGCCACTTTCCTCTCATCCCTCTGGGAGAGCTTTCTGGACAGGTCCTGTC 360
 Db 301 caactggagccacttgcctctcctcctcctcctcctcctcctcctcctcctcctcctc 360
 QY 361 CTCCTTGGGGCCCTGCAGAGCCCTCTTGGAAACCCAGCTTCTCCTCACAGGCGAGGACCA 420
 Db 361 ctcccttggggccctgcagagcctccttggaaacccagcttctcctcacagggcagaccaca 420
 QY 421 GCTCACAGGATCCCAATGCCATCTTCTCCTGAGCTTCCACACCTGTCTCCGAGGAAGGTG 480
 Db 421 gctcacaggatcccaatgccatcttctcctcctcctcctcctcctcctcctcctcctc 480
 QY 481 CGTTTCTCTGATGCTTGTAGAGGTCCTCACCTCTGCGTCAAGGGGTGGC 528
 Db 481 cgtttctctgatgctgttaggaggtccaccctctcctcctcctcctcctcctcctcctc 528

RESULT 13
 AAT47958
 ID AAT47958 standard; cDNA to mRNA; 1086 BP.
 XX
 AC AAT47958;
 XX
 DT 03-APR-1997 (first entry)
 XX
 DE Human thrombopoietin (1-332) coding sequence.
 XX
 KW TPO; recombinant; thrombopoietin; storage; reduced adsorption;
 KW container wall; additive; ds.
 XX
 OS Homo sapiens.
 XX
 Key Location/Qualifiers
 FT 25..1086
 FT CDS /*tag= a
 FT /note= "immature protein"
 FT sig_peptide 25..87
 FT /*tag= b
 FT mat_peptide 88..1083
 FT /*tag= c
 FT /note= "mature protein residues 1-332"
 XX

PN WO9628181-A1.
 XX 19-SEP-1996.
 PD 14-MAR-1996; 96WO-JP00635.
 PF 15-MAR-1995; 95JP-0056248.
 PR (KIRI) KIRIN BREWERY KK.
 PA
 XX Otsuki N;
 PI
 XX WPI; 1996-433541/43.
 DR P-PSDB; AAW09314.
 XX
 XX Stable thrombopoietin compositions contains TPO and an additive -
 PT does not lose thrombopoietin activity on storage in a container
 PT
 XX Example 1; Page 26-28; 38pp; Japanese.
 PS
 XX Stable thrombopoietin (TPO) compsns. containing TPO and an additive,
 CC pref. human serum albumin and/or gelatine, etc., are claimed. The
 CC compsns. do not lose TPO activity on storage in a container as the
 CC compsn. has reduced adsorption onto the container walls. The present
 CC sequence encodes recombinant immature human TPO (1-332).
 XX
 XX Sequence 1086 BP; 229 A; 386 C; 243 G; 228 T; 0 other;

Query Match 48.3%; Score 523.2; DB 17; Length 1086;
 Best Local Similarity 99.4%; Pred. No. 9.8e-116;
 Matches 525; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGGAGCTGACTGAATGCTCCCTCGTGTATGCTTCTCTTAAGGCTAACGCTG 60
 Db 25 atggagctgactgaatgctccctcggtgtatgcttctcttaaggtcaagctaacgctg 84
 QY 61 TCACAGCCGGCTCCTCGTGTGTGACCTCCGAGTCTCAGTAAACTGCTTGTGACTCC 120
 Db 85 tccagccggctcctcctggtgtgaccccgagtcctcagtaactgcttgactcc 144
 QY 121 CATGTCTTCACAGCAGCTGAGCCAGTGCCTGAGAGTTTACCCCTTTGCTTACACCTGTC 180
 Db 145 catgtcttcacagcagctgagccagtgccagaggttcacccctttgcttacacctgtc 204
 QY 181 CTGCTGCTGCTGTGGAGCTTAGCTTGGGAGATGGAAACCCAGATGGAGGACCAAG 240
 Db 205 ctgctgctgctgtggagcttagcttgggagatggaaaaccagatggagagaccag 264
 QY 241 GCACAGGACATTTCTGGAGCAGTGACCTTCTGCTGGAGGAGTGATGGCAGCACGGGA 300
 Db 265 gcacagacattctggagcagtgacccttctgctgagagagtatggcagcagggga 324
 QY 301 CAACTGGAGCCACTTGCCTCTCATCTCCTGCTGGGAGAGTTTCTGACAGGTCCTGTC 360
 Db 325 caactggagccacttgcctctcatcctcctcctggtggcagcttcttgagacaggtccgtc 384
 QY 361 CTCTTGGGGCCCTGACAGGCTCCTTGGAAACCCAGCTTCTCCACAGGGCAGGACACA 420
 Db 385 ctcttggggccctgacagacctcttggaaaccagcttctccacagggagaccaca 444
 QY 421 GCTCACAGGATCCCAATGCCATCTTCTGAGCTTCCACACCTGCTCCGAGGAAAGTG 480
 Db 445 gctcacaggatcccaatgccatcttctgagcttccacacctgctccgaggaaggtg 504
 QY 481 CGTTTCTGTATGCTTGTAGAGGGTCCACCTCTGCTGCTGAGGGTGGC 528
 Db 505 cgttctgtatgcttgttaggaggtccacctctctgctgctgagggggcc 552

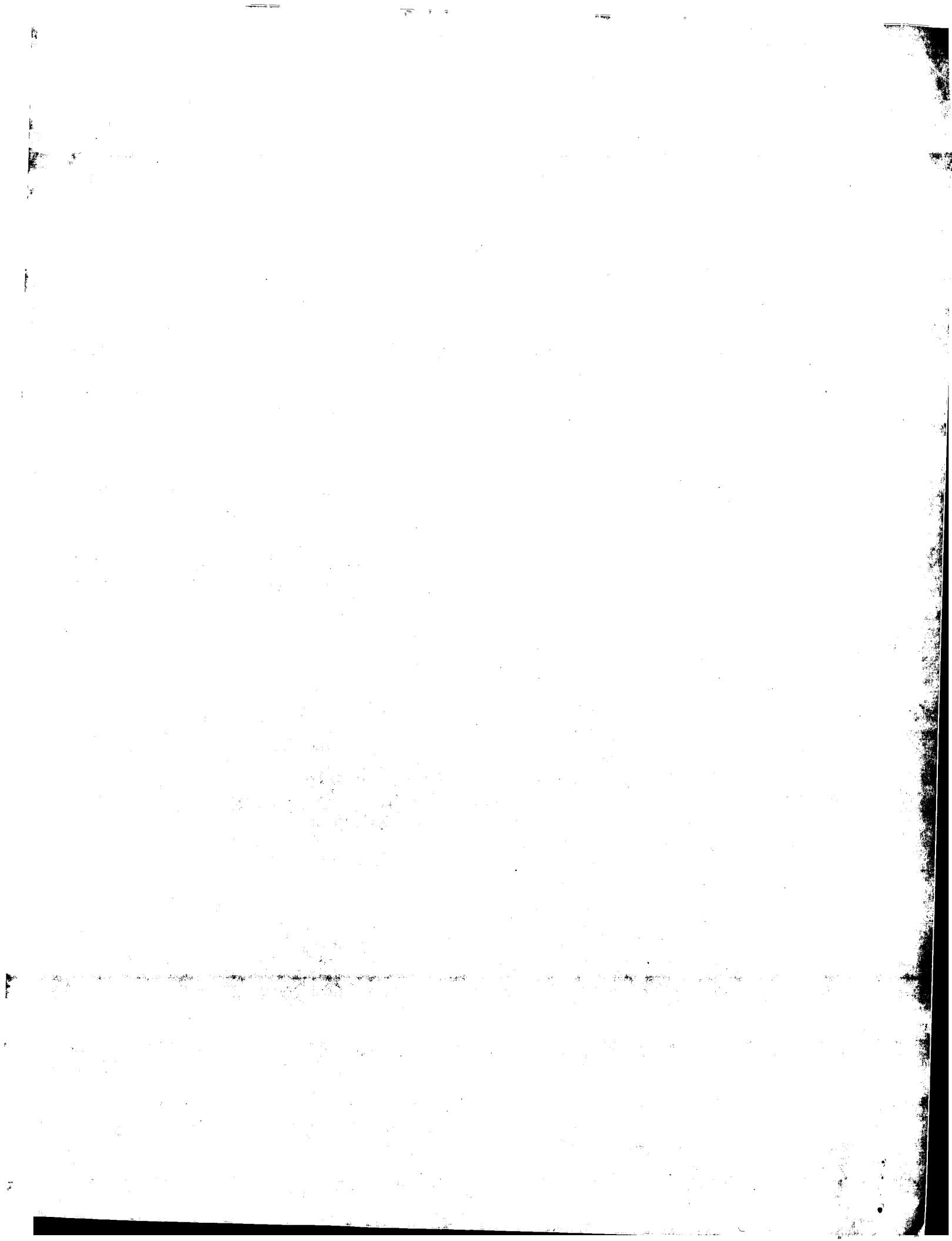
RESULT 14

AAQ99704
 ID AAQ99704 standard; cDNA; 1342 BP.

XX AAQ99704;
 AC 01-FEB-1996 (first entry)
 DT
 XX Human MGDF-1 and MGDF-2.
 DE
 XX Mpl ligand; mammalian megakaryocyte growth promoting factor;
 KW platelet producing factor; ss.
 KW
 XX Homo sapiens.
 OS
 XX Key Location/Qualifiers
 FH 36..1097
 FT CDS /*tag= a
 FT
 XX EP675201-A.
 PN
 XX 04-OCT-1995.
 PD
 XX 30-MAR-1995; 95EP-0104711.
 PF
 XX 30-NOV-1994; 94US-0347780.
 PR 31-MAR-1994; 94US-0221768.
 PR 31-MAY-1994; 94US-0252628.
 PR 12-OCT-1994; 94US-0321488.
 PR
 XX (AMGE-) AMGEN INC.
 PA
 XX Bartley TD, Bogenberger JM, Bosselman RA, Hunt P;
 PI Kinstler OB, Samal BB;
 PI
 XX WPI: 1995-338287/44.
 DR P-PSDB; AAR80821, AAR80822, AAR80823, AAR80824, AAR80825, AAR80826.
 DR
 XX New human megakaryocyte growth and development factor - used to increase
 PT the number of megakaryocytes or platelets in patients, e.g. for treating
 PT thrombocytopenia
 PT
 XX Claim 17; Fig 11; 98pp; English.
 PS
 XX Mpl ligand is synonymous with MGDF. Mpl ligands are mammalian
 CC megakaryocyte growth promoting and/or platelet producing factors.
 CC MGDF-2 is a truncated portion of MGDF-1. It is hypothesized that
 CC human MGDF is expressed in vivo as a substantially inactive or less
 CC active precursor polypeptide that contains variable C-terminal AAs.
 CC Upon cleavage of the C-terminal AAs (as well as the signal peptide),
 CC the processed form(s) of the molecule retain activity or become more
 CC active. It is believed that MGDF-1 may require processing in order to
 CC exhibit its activity. The fact that a truncated form MGDF-1 (i.e.
 CC MGDF-2) is active supports this hypothesis. Various active molecules
 CC that may result from truncations of the sequence set forth as
 CC MGDF-1 (AAR80824) are given in AAR80822, AAR80823, & AAR80825.
 CC is the signal peptide. The preferred truncation variants of MGDF-1
 CC are any of those that have C-terminal truncations from AA 173-353
 CC (along with cleavage of the signal peptide). The signal peptide
 CC may have 23 AAs.
 XX
 XX Sequence 1342 BP; 299 A; 454 C; 293 G; 296 T; 0 other;

Query Match 48.3%; Score 523.2; DB 16; Length 1342;
 Best Local Similarity 99.4%; Pred. No. 1e-115;
 Matches 525; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGGAGCTGACTGAATGCTTCTCGTGGTCTATGCTTCTTAAGGCTAACGCTG 60
 Db 36 atggagctgactgaatgcttctcctcggtgctatgcttctcttaaggtcaagctaacgctg 95
 QY 61 TCACAGCCGGCTCCTCCTGCTTGTGACCTCCGAGTCTCAGTAACTGCTTGTGACTCC 120
 Db 96 tccagccggctcctcctgcttggactcccgagtcctcagtaactgcttctgactcc 155



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 5, 2002, 09:47:00 ; Search time 1909.33 seconds

(without alignments)
6095.163 Million cell updates/sec

Title: US-09-680-514-6

Perfect score: 1083

Sequence: 1 ATGGAGCTGACTGAATTGCT.....TACGCCACTTCCCGAGCCC 1083

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: em_estfun:*
2: em_esthum:*
3: em_estin:*
4: em_estom:*
5: em_estpl:*
6: em_estro:*
7: em_estro:*
8: em_estov:*
9: em_hic:*
10: gb_est1:*
11: gb_est2:*
12: gb_hic:*
13: gb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pin:*
18: em_gss_pro:*
19: em_gss_rod:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	320.4	29.6	973	BI411128	BI411128 602962472
2	312.8	28.9	598	11	BG548320
3	301.6	27.8	900	11	BI328788
4	300.6	27.8	507	10	BE485194
5	300.6	27.8	633	10	AI314551
6	299.6	27.7	609	10	AA538257
7	285.2	26.3	878	11	BF789308
8	227.8	21.0	464	10	BE754587
9	135.8	12.5	327	11	BF848766
10	135.2	12.5	449	13	AZ025993
11	109.2	10.1	419	13	AZ352219
12	86.4	8.0	708	13	AZ792988

13	77.8	7.2	752	13	AQ888740
14	65.6	6.1	589	10	BE377769
15	55.6	5.1	925	13	CNS0091P
16	50.2	4.6	513	10	BE480590
17	49	4.5	925	13	CNS0091P
18	48.8	4.5	617	10	AI981598
19	43.2	4.0	461	10	BE598945
20	43.2	4.0	559	11	BG556176
21	43.2	4.0	976	11	BG823538
22	42.6	3.9	452	10	AA155632
23	42.6	3.9	899	13	CNS015Y1
24	42.4	3.9	1203	13	CNS015Y4
25	42.2	3.9	598	10	AA115932
26	41.4	3.8	421	10	AA058743
27	41.4	3.8	478	11	BI345642
28	40.8	3.8	545	11	BF868325
29	40.6	3.7	447	11	BG604563
30	40.6	3.7	845	10	AL572931
31	40.6	3.7	932	13	CNS0072Q
32	40.6	3.7	1003	11	BI410408
33	40.6	3.7	1101	13	CNS05709
34	40.4	3.7	458	10	AW273202
35	40.4	3.7	485	10	BE463718
36	40.2	3.7	864	13	CNS022LE
37	39.8	3.7	387	10	AW193322
38	39.8	3.7	427	10	BE048584
39	39.8	3.7	432	10	BE175624
40	39.8	3.7	441	10	AI818468
41	39.8	3.7	444	10	AW072844
42	39.8	3.7	450	10	AW273147
43	39.8	3.7	453	10	AW337946
44	39.8	3.7	492	13	AQ686347
45	39.8	3.7	505	10	AI952164

ALIGNMENTS

RESULT 1

BI411128
LOCUS BI411128 973 bp mRNA EST 14-AUG-2001
DEFINITION 602962472F1 NCI_CGAP_Lu33 Mus musculus cDNA clone IMAGE:5118185 5', mRNA sequence.
ACCESSION BI411128 GI:15172051
VERSION BI411128.1
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 973)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA sequencing by: Incyte Genomics, Inc.
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LLAM11287 row: O column: 18
High quality sequence start: 32
High quality sequence stop: 912.

FEATURES

Location/Qualifiers
1..973
/organism="Mus musculus"
/strain="CZECH II"
/db_xref="taxon:10090"
/clone="IMAGE:5118185"

```
/clone_lib="NCI_CGAP_Lu33"
/tissue_type="pooled lung tumors"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: lung; Vector: pT73D-Pac (Pharmacia) with a
modified polylinker; Site_1: NotI; Site_2: EcoRI; 1st
strand cDNA was prepared from mRNA obtained from pooled
lung tumors with a Not I - oligo(dT) primer [5'
TGTACCAATCTGAAGTGGAGCGCGCTCTGTGTTTTTTTTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not
I and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT 192 a 295 c 254 g 232 t
ORIGIN
Query Match 29.6%; Score 320.4; DB 11; Length 973;
Best Local Similarity 77.1%; Pred. No. 1.1e-62;
Matches 390; Conservative 0; Mismatches 116; Indels 0; Gaps 0;
QY 578 CCTCGAGCTTACCACAGAGCTTCCTTTAAAGCTTAGACAGAGTGGAGAGATCCAGG 637
Dy 91 CCCTGCTCTGCCCCGAAGCTTCCTGCTTAAGTCCCTGGAGCAAGTGGAGAGATCCAGG 150
QY 638 GCGATGGCGAGCGCTCCAGGAGAGCTGTGTGCCACCTACAAAGCTGTGCCACCCCGAGG 697
Dy 151 CCACGGCTCGGTGCTGCTGAGCAGTGTGTGCCACCTACAAAGCTGTGTACCCCGAGG 210
QY 698 AGCTGGTGTGCTGGGACACTCTGTGGGCATCCCTGGGGTCCCTGGAGCAGCTGCCCA 757
Dy 211 AGCTGGTGTGCTGGGCACTCTGTGGGATCCGGAAGGCTTCCTGAGTGGCTGCTTA 270
QY 758 GCCAGGCGCTGACGCTGGCAGCTGCTGTGAGCAACTCCATAGCGGCTTTCTCTACC 817
Dy 271 GCCAGGCGCTGACGAGACAGCTGCTTAAGCAGCTCCATAGCGGCTTCGCTCTACC 330
QY 818 AGGGGCTCTGACGCGCTGGAAGGATCTCCCGGAGTTGGTTCACCTTGGACACAC 877
Dy 331 AAGTCTCTGCTGAGGCTCTATCGGATATTTCCCTGCTGCGCCCGCCACCTTGGACTTCG 390
QY 878 TGCAGCTGGAGCTGGCGAGCTTTGCCACCACTATCTGGCAGCAGATGGAAGAACTGGGA 937
Dy 391 TTCAGCTGGATGTGCCAACTTTGCCACCACTATCTGGCAGCAGATGGAAGAACTAGGGG 450
QY 938 TGGCCCTGCTGACGCGCCACCGAGGTCGCATCGCGGCTTCGCTTCTGCTTCCAGC 997
Dy 451 TGGCCCTGCTGACGCGCCACCGAGGTCGCATCGCGGCTTCGCTTCTGCTTCCAGC 510
QY 998 GCGGGCAGGAGGCTCTAGTTGCTTCCATCTGCAGAGCTTCCCTGGAGGTGCTGCTACC 1057
Dy 511 GCGGGCAGGAGGCTCTAGTTGCTTCCATCTGCAGAGCTTCCCTGGAGGTGCTGCTACC 570
QY 1058 GCGTCTACGCCACCTTGGCCAGCC 1083
Dy 571 TTGCTCTGCACCACTTGGCCTAGACC 596
RESULT 2
BG548320 598 bp mRNA EST 04-APR-2001
LOCUS 602575289F1 NIH_MGC_77 Homo sapiens cDNA clone IMAGE:4703159 5',
DEFINITION mRNA sequence.
ACCESSION BG548320
VERSION BG548320.1 GI:13546985
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 598)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
```

```
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLC1541 row: b column: 24
High quality sequence stop: 597.
Location/Qualifiers
1..598
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NIH_MGC_77"
/lab_host="DH10B (T1 phage-resistant)"
/notes="Organ: lung; Vector: pDNR-LIB (Clontech); Site_1:
SfiI (ggccattatgcc); Site_2: SfiI (ggccattatgcc); 5' and
3' adaptors were used in cloning as follows: 5' adaptor
sequence: 5'-CACGCCCTATGGC-3' and 3' adaptor sequence:
5'-ATTCTAGAGCGGCGGCGGACATG-dT(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.9
kb (range 0.5-4.0 kb). 12/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA). Note: this is a NIH_MGC Library."
BASE COUNT 113 a 202 c 163 g 120 t
ORIGIN
Query Match 28.9%; Score 312.8; DB 11; Length 598;
Best Local Similarity 99.4%; Pred. No. 5.1e-61;
Matches 314; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 768 GCAGCTGGCAGGCTGTGTAGCCAACTCCATAGCGGCTTTCTCTACAGGGGCTCCT 827
Dy 185 GAAGCTGGCAGGCTGTGTAGCCAACTCCATAGCGGCTTTCTCTACAGGGGCTCCT 244
QY 828 GCAGGCGCTTGGAGGATCTCCCGGAGTTGGTCCACCTTGGACACACTGCAGCTGGA 887
Dy 245 GCAGGCGCTTGGAGGATCTCCCGGAGTTGGTCCACCTTGGACACACTGCAGCTGGA 304
QY 888 CGTGGCGGCTTGGCACCACCATCTGGCAGCAGATGGAAGAACTGGGAATGGCCCTGC 947
Dy 305 CGTGGCGGCTTGGCACCACCATCTGGCAGCAGATGGAAGAACTGGGAATGGCCCTGC 364
QY 948 CTGCAGCGCCACCGAGGTCGCATCGCGGCTTCGCTTCTGCTTCCAGCGCGGCGAGG 1007
Dy 365 CTGCAGCGCCACCGAGGTCGCATCGCGGCTTCGCTTCTGCTTCCAGCGCGGCGAGG 424
QY 1008 AGGGTCTCTAGTTGCTTCCCATCTGCAGAGCTTCTCTGGAGGTGTCGACCGCTTACG 1067
Dy 425 AGGGTCTCTAGTTGCTTCCCATCTGCAGAGCTTCTCTGGAGGTGTCGACCGCTTACG 484
QY 1068 CCACCTTGGCCAGCCC 1083
Dy 485 CCACCTTGGCCAGCCC 500
RESULT 3
BG1328788 900 bp mRNA EST 30-JUL-2001
LOCUS 602984776F1 NCI_CGAP_L19 Mus musculus cDNA clone IMAGE:5137578 5',
DEFINITION mRNA sequence.
ACCESSION BG1328788
VERSION BG1328788.1 GI:15013445
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

Email: causerps1.baird.usda.gov
Single pass sequencing. Bases called and alt_trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -minmatch 12 options.

```

vv.900904.e: vector identified by cross_match with the minuscore 1
and -minmatch 12 options.
PCR primers
FORWARD: AGGAAACAGCTATGACCAT
BACKWARD: GTTTTCCCAGTCACCAAG
Plate: 134 row: J column: 17
Seq primer: ATTTAGTGACACTATAG.

```

```

BASE COUNT
ORIGIN
97 a 170 c 158 g 82 t
/ab_xref="taxon:9913
/clone_lib="BAC 5BOV"
/tissue_type="pooled"
/tiss_host="DH10b"
/note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
Library made from pooled mRNA isolated from mammary
tissues at eight physiological, developmental, and disease
states."

```

Query Match	Score	DB 10;	Length	507;
Best Local Similarity	84.1%;	Pred. No. 2.9e-58;		
Matches 339; Conservative	0;	Mismatches 64; Indels	0;	Gaps
0;				
QY	666	GTGTGCCACCTACAAGCTGTGCCACCCCGAGGAGCTGGTGTCTGCTCGACACACTCTCTGGG	725	
Db	104	GTGTGCCGCCACCAAGCTGTGCCACCCCGAGGAGCTGATGCTGTCTCAGGCACACTCTCTGGG	163	
QY	726	CATGCCCTGGGCTCCCTGTGACAGCTGCCCGACGAGGCCCTGCAGCTGGCAGGCTGCTT	785	
Db	164	CATCCCCCAGGCTCCCCTAAGCAGCTGCTCCAGGCAGTCCCTGCAGCTGCAGACTGCCT	223	
QY	786	GAGCAACTCCATAGCGGCCCTTTTCTCTACCAAGGGGCTCCTGCAGGCCCTTGGAAAGGAT	845	
Db	224	GAACCAACTACAGCGCGGCCCTTCTCTACCAAGGGGCTCCTGCAGGCCCTTGGCGGGCAT	283	
QY	846	CTCCCCCGAGTTGGTCCCCACCTTGGACACACTGCAGCTGGACGTGCGCGACTTTGGCAC	905	
Db	284	CTCCCCAGAGTGGCCCCCACCTTGGACACACTGCAGCTGGACGTCTACTGACTTGGCAC	343	
QY	906	CACCATCTGCAGCAGATGGAAGACTGGAAATGGCCCCCTGCCCTGCAGCCCCACCCAGGG	965	
Db	344	GAACATCTGGCTGCAGATGAGGACCTGGGGGGGGCCCCGCTGTGCAGCCCCACCCAGGG	403	
QY	966	TGCCATCGCGGCTTCGCCCTCTGCTTTCCAGCGCGCGGACGAGGGGTCTCTAGTTGGCTC	1025	
Db	404	CGCCATGCCACCTTCACTTCAGCCTTCCAAACGACGAGCAGGAGGGTCTCTGGTTGCTTC	463	
QY	1026	CCATCTGCAGAGCTTCTCGAGGTGTGCTACCGCGTTCTACG	1068	
Db	464	CCAGCTGCATCTGTTTCTCTGGAGCTGGCATCCGCTGGGCTCGCG	506	

Db	QY	Db	RESULT	5
404	CGCCATGCCGACCTTCACTTACGCTTCCAAACGACAGCAGGAGGGTCTGGTTGCTTC	463		
1026	CCATCTCGAGAGCTTCCTGGAGGTGTCGTACCGCGTTCTACGC	1068		
464	CCAGCTGCATCGTTTCTCTGGAGCTGGCATACCGTGGCCTGCGC	506		

Db	358	CACCTCCTTCACAGCCGACTGAGTCAAGTGTCCGAGCTGACCCCTTGTCTATCCCTGTT	417
Qy	181	CTGCTGCTGCTGTGGACTTTAGCTTTGGGAGATGGAACCCAGATGAGGACGACCAAG	240
Db	418	CTGCTGCTGCTGTGGACTTTAGCTTTGGGAGATGGAACCCAGACGACGACGACGACAAAG	477
Qy	241	GCACAGGACATCTCGGAGAGCAGTGACCTTCTCTGGAGGAGTGTATGCGACGACGGGA	300
Db	478	GCACAGGACATCTTAGGGGAGTGTCCCTTCTACTTGGAGGAGTGTATGCGACGACGAGGA	537
Qy	301	CAACTGGAGCCACTTGTCCCTCTCATCCCTCTGGGCGACTTTCTTGGACAGGTCCTCTCTC	360
Db	538	CAGTTTGAACCCCTCTCCCTCTCTCATCCCTCTGGACAGCTNCTTGGCAGGTTTCGCCCTC	597
Qy	361	CTCTTGGGGCCCTGCAGAGCCCTCTTGGAAACCCAG	396
Db	598	CTCTTGGGGCCCTGCAGGGCCCTCTAGGAACCCAG	633
RESULT	6		
AA538257			
LOCUS	AA538257	609 bp	mRNA
DEFINITION	vJ03a07.r1 Barstead mouse pooled organs MPLRB4 Mus musculus cDNA		29-JUL-1997
	clone IMAGE:920628 5' similar to gb:L34169 Mus musculus		
	thrombopoietin mRNA, complete cds (MOUSE);		mRNA sequence.
ACCESSION	AA538257		
VERSION	AA538257.1	GI:2284250	
KEYWORDS	house mouse.		
SOURCE	Mus musculus		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
REFERENCE	1 (bases 1 to 609)		
AUTHORS	Maria,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,		
	Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,		
	Schellenberg,K., Steptoe,K., Tan,F., Underwood,K., Moore,B.,		
	Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and		
	Waterson,R.		
TITLE	The WashU-HMI Mouse EST Project		
JOURNAL	Unpublished (1996)		
COMMENT	Contact: Marra M/Mouse EST Project		
	WashU-HMI Mouse EST Project		
	Washington University School of MedicineP		
	4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108		
	Tel: 314 286 1800		
	Fax: 314 286 1810		
	Email: mouseest@watson.wustl.edu		
	This clone is available royalty-free through LLNL; contact the		
	IMAGE Consortium (info@image.llnl.gov) for further information.		
	MGI:532844		
	Seq primer: -28m13 rev2 ET from Amersham		
	High quality sequence stop: 492.		
FEATURES	Location/Qualifiers		
source	1..609		
	/organism="Mus musculus"		
	/strain="FVB/N"		
	/db_xref="taxon:10090"		
	/clone_lib="IMAGE:920628"		
	/clone_lib="Barstead mouse pooled organs MPLRB4"		
	/sex="mixed"		
	/tissue_type="pooled organs"		
	/dev_stage="7 day"		
	/lab_host="DH10B"		
	/note="organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with		
	a modified polyLinker; Site_1: EcoRI; Site_2: NotI; 1st		
	strand cDNA was primed with a Not I - oligo(dT) primer. [5		
	TGTTCAGTAATCTGAAGTGGAGCGCGCCCTTTTTTTTTTTTTTTTTTTTTTTTTTTT		
	3']; double-stranded cDNA was ligated to Eco RI adaptors		
	[GTTGATTCGGTACC], digested with Not I and cloned into		
	the Not I and Eco RI sites of the modified pT7T3 vector.		
	Library constructed by Bob Barstead."		
BASE COUNT	126 a 204 c 157 g 122 t		
ORIGIN			

AI314551	633 bp	mRNA	EST	17-DEC-1998
u49c10.y1	sugano mouse liver mlia	Mus musculus	cDNA clone	
IMAGE:1923186	5' similar to gb:U34169	Mus musculus	thrombopoietin mRNA, complete cds (MOUSE);	mRNA sequence.
AI314551				
AI314551.1	GI:4029818			
EST.				
Mus musculus				
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus;				
1 (bases 1 to 633)				
Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wyllie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.				
The WashU-HMI Mouse EST Project				
Unpublished (1996)				
Contact: Marra M/Mouse EST Project				
WashU-HMI Mouse EST Project				
Washington University School				
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108				
Tel: 314 286 1800				
Fax: 314 286 1810				
Email: mousest@wustl.wustl.edu				
This clone is available royalty-free through LLMU ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.				
MGI:979478				
Seq primer: custom primer used				
High quality sequence stop: 505.				
Location/Qualifiers				
1. 633				
/organism="Mus musculus"				
/strain="C57BL"				
/db_xref="taxon:10090"				
/clone="IMAGE:1923186"				
/clone_lib="Sugano mouse liver mlia"				
/sex="female"				
/dev_stage="adult"				
/lab_host="DH10B"				
/note="Organ: liver; Vector: pME18S-FL3; Site_1: DraIII (CACTGCTG); Site_2: DraII (CACCATGTG); 1st strand cDNA was primed with an oligo(dT) primer [ATGTGGCCCTTTTCTTTTCTTTT]; double-stranded cDNA was ligated to a DraIII adaptor [TGTGGCCCTACTGG], digested and cloned into distinct draIII sites of the pME18S-FL3 vector (5' site CACTGCTG, 3' site CACCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTCTGCTCTCTAAAGCTGCG and 3' end primer CGACCTGACGCTGAGACAC."				
125 a	210 c	166 g	131 t	1 others
BASE COUNT				
ORIGIN				
Query Match	27.8%;	Score 300.6;	DB 10;	Length 633;
Best Local Similarity	84.8%;	Prod. No. 3e-58;		
Matches 336;	Conservative	0;	Mismatches 60;	Indels 0; Gaps 0;
QY	1	ATGAGCTGACGTGAATTGCTCTCGTGTGTCATGCTCTCTCACTGCAAGGCTAACGCTG	60	
Db	238	ATGGAGCTGACGTGATTGCTCTCGGCGCATCTCTTGTGAGTGGCAGACTACTCTG	297	
QY	61	TCCAGCCCGCCTCTCTGTTGTGACCTCCGAGTCTCTCAGTAAACTGCTTCGTGACTCC	120	
Db	298	TCCAGCCCGTAGTCTCTGCTGTGACCCAGACTCCTAAATAAACTGCTGCTGACTCC	357	
QY	121	CATGCTCTTCACAGCAGACTGAGCCAGTCCAGAGGTTTCACCCCTTTCCTACACTGTC	180	

Query Match 27.7%; Score 299.6; DB 10; Length 609;
Best Local Similarity 81.6%; Pred. No. 5.1e-58;
Matches 364; Conservative 0; Mismatches 69; Indels 13; Gaps 1;

QY 1 ATGGAGCTGACTGAATTTGCTCTCGTGGTCAATGCTTCTCTAACTGCAAGGCTTAACGCTG 60
DB 177 ATGGAGCTGACTGATTTGCTCTCGTGGGCCATGCTTCTTGCAGTGGCAAGACTAACTCTG 236
QY 61 TCAGAGCCGCTCTCTCTGCTGTGACCTCCGAGTCCAGTAACTGCTTGTGACTCC 120
DB 237 TCAGAGCCGCTAGTCTCTGCTGTGACCCAGACTCTCTAAATAAATGCTGCGTGAATCC 296
QY 121 CATGTCTCTCACAGCAGTGAAGCAGTGCAGGCTTCCACCTTTGCTTACACCTGTC 180
DB 297 CACTCTCTCACAGCAGTGAAGTGTGCTGCTGCTGCTGCTGCTTGTCTATCCCTGTT 356
QY 181 CTGCTGCTGCTGTGACTTTAGCTTGGGAGAAATGGAACCCAGATGAGGAGACCAAG 240
DB 357 CTGCTGCTGCTGTGACTTTAGCTTGGGAGAAATGGAACCCAGACGGAACAGCAAG 416
QY 241 GCACAGACATTTGAGGAGCAGTGAAGCTTCTGCTGAGGAGTGTGAGCAGCAGCGGA 300
DB 417 GCACAGACATTTAGGAGCAGTGTCCCTTCTACTGAGGAGTGTGAGCAGCAGGAGGA 476
QY 301 CAATGGGAGCCCTGAGAGCCCTCTTGGAAACCCAGCTTCTGAGCAGGTCGCTC 360
DB 477 CAGTTGGAACCTCTGCTCTCTATCCCTCTGAGCAGCTTCTGAGCAGGTTGCGCTC 536
QY 361 CTCTTGGGCGCTGAGAGCCCTCTTGGAAACCCAGCTTCTGAGCAGGAGCAGCACA 420
DB 537 CTCTTGGGCGCTGAGAGCCCTCTTGGAAACCCAGCTTCTGAGCAGGAGCAGCACA 583
QY 421 GCTCAGAGATCCCAATGCTCTT 446
DB 584 GCTCAGAGATCCCAATGCTCTT 609

RESULT 7
BF789308 878 bp mRNA EST 12-JAN-2001
LOCUS 602104729F1 NCI_CGAP_Kid14 Mus musculus cDNA clone IMAGE:422927
DEFINITION 5', mRNA sequence.
ACCESSION BF789308
VERSION BF789308.1 GI:12094344
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 878)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
JOURNAL NIH-MGC <http://mgc.nci.nih.gov/>.
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cga@rs-research.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
• Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM9810 row: 1 column: 08
High quality sequence stop: 628.
Location/Qualifiers
1. 878
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:422927"
/clone_lib="NCI_CGAP_Kid14"

FEATURES
Source

/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: Kidney; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.75 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library. 1"

BASE COUNT 203 a 281 c 224 g 169 t 1 others
ORIGIN

Query Match 26.3%; Score 285.2; DB 11; Length 878;
Best Local Similarity 84.7%; Pred. No. 1e-54;
Matches 320; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 1 ATGGAGCTGACTGAATTTGCTCTCGTGGTCAATGCTTCTCTAACTGCAAGGCTTAACGCTG 60
DB 238 ATGGAGCTGACTGATTTGCTCTCGTGGGCCATGCTTCTTGCAGTGGCAAGACTAACTCTG 297
QY 61 TCAGAGCCGCTCTCTCTGCTGTGACCTCCGAGTCCAGTAACTGCTTGTGACTCC 120
DB 298 TCAGAGCCGCTAGTCTCTGCTGTGACCCAGACTCTCTAAATAAATGCTGCGTGAATCC 357
QY 121 CATGTCTCTCACAGCAGTGAAGCAGTGCAGGCTTCCACCTTTGCTTACACCTGTC 180
DB 358 CACTCTCTCACAGCAGTGAAGTGTGCTGCTGCTGCTGCTTGTCTATCCCTGTT 417
QY 181 CTGCTGCTGCTGTGACTTTAGCTTGGGAGAAATGGAACCCAGATGAGGAGACCAAG 240
DB 418 CTGCTGCTGCTGTGACTTTAGCTTGGGAGAAATGGAACCCAGACGGAACAGCAAG 477
QY 241 GCACAGACATTTGAGGAGCAGTGAAGCTTCTGCTGAGGAGTGTGAGCAGCAGCGGA 300
DB 478 GCACAGACATTTAGGAGCAGTGTCCCTTCTACTGAGGAGTGTGAGCAGCAGGAGGA 537
QY 301 CAATGGGAGCCCTGAGAGCCCTCTTGGAAACCCAGCTTCTGAGCAGGAGCAGCACA 360
DB 538 CAGTTGGAACCTCTGCTCTCTATCCCTCTGAGCAGCTTCTGAGCAGGTTGCGCTC 597
QY 361 CTCTTGGGCGCTGAGAG 378
DB 598 CTCTTGGGCGCTGAGAG 615

RESULT 8
BF754587 464 bp mRNA EST 25-APR-2001
LOCUS 208137 MARC 2BOV Bos taurus cDNA 5', mRNA sequence.
DEFINITION BF754587
ACCESSION BF754587
VERSION BF754587.1 GI:10168579
KEYWORDS EST.
SOURCE COW.
ORGANISM Bos taurus
REFERENCE 1 (bases 1 to 464)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
JOURNAL Bovidae; Bovinae; Bos.
COMMENT Smith, T. P. L., Grosse, W. M., Freking, B. A., Roberts, A. J., Stone, R. T.,
Casas, E., Wray, J. E., White, J., Cho, J., Fahrenkrug, S. C., Bennett,
G. L., Heaton, M. P., Laegreid, W. W., Rohrer, G. A., Chitko-McKown, C. G.,
Perle, G., Holt, I., Karamycheva, S., Liang, F., Quackenbush, J. and
Keefe, J. W.
Sequence evaluation of four pooled-tissue normalized bovine cDNA
libraries and construction of a gene index for cattle
Genome Res. 11 (4), 626-630 (2001)
JOURNAL 21180013
MEDLINE
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt_trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -minmatch 12 options.

PCR Primers
 FORWARD: AGGAACAGCTATGACCAT
 BACKWARD: GTTTCCAGTCACGACG
 Plate: 54 row: J column: 23
 Seq primer: ATTAGGTGACATATAG.
 Location/Qualifiers
 1. .464

FEATURES

source
 /organism="Bos taurus"
 /db_xref="taxon:9913"
 /clone_lib="MARC 2BOV"
 /tissue_type="pooled"
 /lab_host="DH10B"

/note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
 Library made from pooled tissue from testis, thymus,
 semitendinosus muscle, longissimus muscle, pancreas,
 adrenal, and endometrium."
 97 a 152 c 128 g 87 t

BASE COUNT

ORIGIN
 21.0%; Score 227.8; DB 10; Length 464;
 Best Local Similarity 87.2%; Pred. No. 9.7e-42;
 Matches 266; Conservative 0; Mismatches 27; Indels 12; Gaps 1;

Query Match
 Best Local Similarity 87.2%; Score 227.8; DB 10; Length 464;
 Matches 266; Conservative 0; Mismatches 27; Indels 12; Gaps 1;

Qy 203 GCTTGGAGAAATGGAACCCAGATGGAGGAGACCAAGCAGGACATCTTGGGAGCAG 262
 Db 1 GCTTGGAGAAATGGAACCCAGATGGAGGAGACCAAGCAGGACATCTTGGGAGCAG 60
 Qy 263 TGACCTTCCTGCTGGAGGAGTATGGCAGCAGCGGGGACCACTGGGACCTTGCCTCT 322
 Db 61 CGACCTTCCTGCTGGAGGAGTATGGCAGCAGCGGGGACCACTGGGACCTTGCCTCT 120
 Qy 323 CATCCCTCCTGGGCGCAGCTTCTGGACAGGCTCGCTCTCTCTGGGGCCCTGCAGAGCC 382
 Db 121 CATCCCTCCTGGGCGCAGCTTCTGGGCGCAGCTCGCTCTCTGGGGCCCTGCAGAGCC 180
 Qy 383 TCCTTGAACCCAGCTTCTCCACAGGCGAGGACCAACAGCTCACAGGATGCCATGCCA 442
 Db 181 TCTTAGAACCC-----CAGGGCAGGACCAACAGCTCACAGGATGCCATGCCA 228
 Qy 443 TCTTCTGAGCTTCCCAACACCTGCTCCGAGGAAAGTGCTTCTGATGCTTGTAGGAG 502
 Db 229 TCTTCTGAGCTTCCCAACAGCTGCTCCGAGGAAAGTGCTTCTGATGCTTGTAGGAG 288
 Qy 503 GGTCC 507
 Db 289 GGGCC 293

RESULT 9

BF848766
 LOCUS
 DEFINITION
 QV0-EN0102-081100-458-g06 EN0102 Homo sapiens cDNA, mRNA sequence.
 ACCESSION
 BF848766
 VERSION
 BF848766.1 GI:12235903
 KEYWORDS
 EST.
 SOURCE
 human.

ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1 (bases 1 to 327)
 Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
 Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
 Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,
 M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
 Simpson,A.J.
 Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 JOURNAL
 MEDLINE
 20202663
 COMMENT
 Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?ti=QV0&t2=QV0-EN0102-
 081100-458-g06&t3=2000-11-08&t4=1)
 Seq primer: puc 18 forward
 High quality sequence stop: 327.

FEATURES

source
 1. .327
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="EN0102"
 /dev_stage="Adult"
 /note="Organ: lung_normal; Vector: puc18; Site_1: SmaI;
 Site_2: SmaI; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the puc 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."
 61 a 87 c 103 g 76 t

Query Match
 Best Local Similarity 89.6%; Pred. No. 7.1e-21;
 Matches 146; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 762 GGCCTGCGAGCTGGAGGCTGCTTGGACCAATCCATAGCGGCTTTCTCTACACAGG 821
 Db 139 GGCCTATACACAGGCGCGCTGTTGAGCAATCCATAGCGGCTTTCTCTACACAGG 198
 Qy 822 GTCCTGCGAGCCCTGGAAGGATCTCCCGGAGTTGGTCCACCTTGGACACATGCA 881
 Db 199 GCTCTGCGAGCCCTGGAAGGATCTCCCGGAGTTGGTCCACCTTGGACACATGCA 258
 Qy 882 GCTGACGCTCGCGACTTTCGCCACCACTCTGGCAGCAGATG 924
 Db 259 GCTGACGCTCGCGACTTTCGCCACCACTCTGGCAGCAGATG 301

RESULT 10

AZ025993
 LOCUS
 DEFINITION
 AZ025993
 AZ025993.1 GI:7101377
 VERSION
 AZ025993.1 GI:7101377
 KEYWORDS
 GSS.
 SOURCE
 house mouse.
 ORGANISM
 Mus musculus

REFERENCE
 1 (bases 1 to 449)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 Z hao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S., Akinret
 ,B., Levins,M., McGann,S., Tsegaye,G., Geer,K., Krol,M., de Jong,P.
 and Fraser,C.M.
 Mouse BAC End Sequences from Library RPCI-23
 Unpublished (1999)
 Other GSSs: RPCI-23-343116-TV
 CONTACT: Shaying Zhao

TITLE
 JOURNAL
 COMMENT
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: szhao@tigr.org
 Clones are derived from the mouse BAC library RPCI-23. For BAC

library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (<http://bacpac.med.buffalo.edu/orderingframe.htm>) or from Resea ch Genetics ([info@resgen.com](http://www.tigr.org/tadb/bac_ends/mouse/bac_end_intro.html)). BAC end page: http://www.tigr.org/tadb/bac_ends/mouse/bac_end_intro.html

Plate: 343 row: I column: 16
Seq primer: SP6
Class: BAC ends.
Location/Qualifiers
1. 449

FEATURES

source
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-23-343116"
/clone_lib="RPCI-23"
/sex="Female"
/lab_host="DH10B"

/note="Organ: Kidney/Brain; Vector: pBACe3.6; Site:1: EcoRI; Site:2: EcoRI; Female C57BL/6J mouse kidney and/or brain genomic DNA was isolated and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBACe3.6 vector at the EcoRI sites. The ligation products were transformed into DH10B electrocompetent cells (BRL Life Technologies)."

BASE COUNT 90 a 121 c 108 g 129 t 1 others
ORIGIN

Query Match 12.5%; Score 135.2; DB 13; Length 449;
Best Local Similarity 84.4%; Pred. No. 1e-20;
Matches 152; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY 228 GGAGGAGACCAAGCAGACACATCTGGGAGCAGTGACCTCTGCTGGAGGAGTGAT 287
Db 126 GGAACAGCAGACGACAGACATCTAGGGGAGTGCTTTCTACTGGAGGAGTGAT 185

QY 288 GGCAGCAGGGGACAACTGGGAGCCCACTTGCCTCTCATCCCTCTGGGCGAGCTTTCTGG 347
Db 186 GGCAGCAGGAGACAGTGGAAACCTCTCTCATCCCTCTGGGAGCTTTCTGG 245

QY 348 ACAGGTCCGTCCTCTCTGGGCGCTTGAGAGCCTCTTGGACCCAGCTTCTCCACCA 407
Db 246 GCAGGTTCGCTCTCTTGGGGCGCTTGAGAGCCTCTTGGAGCCAGGTAAGTCCCA 305

RESULT 11

AZ352219/c
LOCUS 419 bp DNA GSS 29-SEP-2000
DEFINITION 1M0090P05R Mouse 10kb plasmid UUGCLM library Mus musculus genomic clone UUGCLM0090P05 R, DNA sequence.

ACCESSION AZ352219
VERSION AZ352219.1 GI:10431456
KEYWORDS GSS.

SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 419)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D., Weiss,R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb

Plasmid inserts

JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT 84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00
Plate: 0090 row: P column: 05
Seq primer: CACACAGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 419.
Location/Qualifiers
1. 419

FEATURES

source
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGCLM0090P05"
/clone_lib="Mouse 10kb plasmid UUGCLM library"
/sex="Male"

/lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory DNA Resource (<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gil4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 100 a 98 c 124 g 97 t
ORIGIN

Query Match 10.1%; Score 109.2; DB 13; Length 419;
Best Local Similarity 79.6%; Pred. No. 8.3e-15;
Matches 129; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 360 CCTCTTGGGCGCTCGAGAGCCTCTTGGAAACCCAGCTTCTCCACAGGCGGACCCAC 419
Db 292 CCTACTCACTGGCTCAGGCGCCATCCTCTGCTCAGCTTCTCTACAGGCGGACCCAC 233

QY 420 AGCTCACAAGGATCCCAATGCCATCTCTTGAGCTTCCACACACTGCTCCGAGAAAGGT 479
Db 232 AGCTCACAAGGACCCCAATGCCCTCTTGTGAGCTTGCAACAACACTGCTCGGGAAAGGT 173

QY 480 GCGTTTCTCATGCTTGTAGAGGGTCCACCCCTCTCGGTCTCAG 521

Db 172 GCCTTCTCTGCTTCTGTAGAGGTCCACCCCTCTGTGTCTCAG 131

RESULT 12

AZ792988
LOCUS 708 bp DNA GSS 16-FEB-2001
DEFINITION 2M0046105F Mouse 10kb plasmid UUGCLM library Mus musculus genomic clone UUGCM2M0046105 F, DNA sequence.

ACCESSION AZ792988
VERSION AZ792988.1 GI:12937676
KEYWORDS GSS.

SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 708)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D., Weiss,R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb

Plasmid inserts

JOURNAL
COMMENT

Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308 Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0046 Row: I Column: 05
Seq primer: CGTTGTAAACGACGCGCCAGT
Class: plasmid ends
High quality sequence stop: 708.

FEATURES

source

1. .708
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0046105"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and 14
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptored DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PWD42 (g1473211419b1AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptored mouse DNA was annealed to
adaptored vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT
ORIGIN

156 a 213 c 147 g 192 t
Query Match 8.0%; Score 86.4; DB 13; Length 708;
Best Local Similarity 73.0%; Pred. No. 1.4e-09;
Matches 111; Conservative 0; Mismatches 41; Indels 0; Gaps 0;
Qy 16 TTGCTCCTCGTGTCATGCTTCTCTCACTGCAAGGCTAACGCTGCCAGCCCGGCTCCT 75
Db 1 TTGCTCCTCGGCGCATGCTTCTTGCACTGGCAAGCTAACTCTGTCCAGCCCGCTAGCT 60
Qy 76 CTTGCTGTGACCTCGGAGTCTCAGTAACTGCTTCGTGACTCCCATGCTCTTCACAGC 135
Db 61 CTTGCTGTGACCCAGACTCCTAAATAAAGTCTGCTGCTGACTCCACCTCTTCACAGC 120
Qy 136 AGACTGAGCCAGTGCCCGAGGTTACCCCTTT 167
Db 121 CGACTGGTGAGCAACCCCAAGCCCGCTGCTTT 152

RESULT 13

A0888740

LOCUS

DEFINITION

A0888740 752 bp DNA GSS 10-NOV-1999

HS_3106_B1_B04_MR CIT Approved Human Genomic Sperm Library D Homo

sapiens genomic clone Plate=3106 Col=7 Row=D, DNA sequence.

ACCESSION

A0888740

VERSION

A0888740.1

KEYWORDS

GSS.

SOURCE

human.

REFERENCE

AUTHORS

Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T., and
Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D., and
Hood, L.

TITLE

JOURNAL

MEDLINE

COMMENT

Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)

Contact: Mahairas GG, Wallace JC, Hood L

High Throughput Sequencing Center

University of Washington

401 Queen Anne Avenue North, Seattle, WA 98109, USA

Tel: (206) 616-3618

Fax: (206) 616-3887

Email: jwallace@u.washington.edu

Clones may be purchased from Research Genetics (info@resgen.com).

BAC end Web Server: http://www.htsc.washington.edu

Plate: 3106 row: D column: 7

Seq primer: M13 Reverse

Class: BAC ends

High quality sequence stop: 752.

FEATURES

source

1. .752
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Plate=3106 Col=7 Row=D"
/clone_lib="CIT Approved Human Genomic Sperm Library D"
/sex="male"
/note="Organ: sperm; Vector: pBelobAC11; BAC Clones in
E-Coli DH10B"

BASE COUNT 181 a 219 c 177 g 175 t

ORIGIN

Query Match 7.2%; Score 77.8; DB 13; Length 752;
Best Local Similarity 68.0%; Pred. No. 1.3e-07;
Matches 123; Conservative 0; Mismatches 57; Indels 1; Gaps 1;
Qy 13 GAATGCTCCTCGTGTCATGCTTCTCTCACTGCAAGGCTAACGCTGTCCAGCCCGGCT 72
Db 509 GAATGCTCCTCGTGTCATGCTTCTCTCACTGCAAGGCTAACGCTGTCCAGCCCGGCT 568
Qy 73 CTTCTGCTGTGACCTCGGAGTCTCTCACTGCAAGGCTAACGCTGTCCAGCCCGGCT 132
Db 569 CTTCTGCTGTGACCTCGGAGTCTCTCACTGCAAGGCTAACGCTGTCCAGCCCGGCT 627
Qy 133 AGCAGACTGAGCCAGTGCAGAGGTTACCCCTTTGCCTACACCTGTCTGCTGCTGCT 192
Db 628 GCAAACTGGTGAGATCTCCACACAGTATCCCTTTATCCGCTTAAGTGTAGACCATAC 687
Qy 193 G 193
Db 688 G 688

RESULT 14

BE377769

LOCUS

DEFINITION

ACCESSION

BE377769

VERSION

BE377769.1

KEYWORDS

EST.

SOURCE

house mouse.

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

AUTHORS

TITLE

JOURNAL

BE377769 589 bp mRNA EST 21-JUL-2000
601229920F1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:3593885 5',
mRNA sequence.

BE377769

BE377769.1

GI:9323134

EST.

house mouse.

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 589)

NIH-MGC http://mgc.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 5, 2002, 10:29:34 ; Search time 2520.15 Seconds
(without alignments)
7167.991 Million cell updates/sec

Title: US-09-680-514-8

Perfect score: 1095

Sequence: 1 ATGGAGCTGACTGAATTGCT.....TACGCCACCTTGCCAGCCC 1095

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_om.*

20: em_or.*

21: em_ov.*

22: em_pat.*

23: em_ph.*

24: em_pl.*

25: em_ro.*

26: em_sts.*

27: em_sy.*

28: em_un.*

29: em_vi.*

30: em_htgo_hum.*

31: em_htgo_inv.*

32: em_htgo_rod.*

33: em_htg_hum.*

34: em_htg_inv.*

35: em_htg_rod.*

36: em_htg_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

8

Result No.	Score	Query Match	Length	DB ID	Description
1	520.4	47.5	525	6	E02574
2	518.8	47.4	555	22	E11965
3	518.8	47.4	1059	6	AR103959 Sequence
4	518.8	47.4	1062	6	AR087133 Sequence
5	518.8	47.4	1062	6	AR087571 Sequence
6	518.8	47.4	1062	6	I49760 Sequence 3
7	518.8	47.4	1062	9	H0MTHROMB
8	518.8	47.4	1086	6	E16669
9	518.8	47.4	1267	6	E12182
10	518.8	47.4	1267	6	E12214
11	518.8	47.4	1342	6	AR008878
12	518.8	47.4	1342	6	AR023468
13	518.8	47.4	1342	6	AR087573
14	518.8	47.4	1342	6	AR087573
15	518.8	47.4	1691	9	D32047
16	518.8	47.4	1721	6	E16668
17	518.8	47.4	1721	22	E11961
18	518.8	47.4	1795	6	I33525
19	518.8	47.4	1795	9	H0MCLMPL
20	517.2	47.2	605	6	AR008879
21	517.2	47.2	605	6	AR087574
22	517.2	47.2	605	6	I85130
23	515.6	47.1	1341	9	HSU11025
24	514	46.9	525	6	E02575
25	514	46.9	1062	9	HSU59493
26	514	46.9	1062	9	HSU59494
27	512.4	46.8	525	6	E02576
28	486.8	44.5	525	6	AR024358
29	486.8	44.5	525	6	AR091731
30	486.8	44.5	525	6	E01731
31	486.8	44.5	525	6	E02573
32	486.8	44.5	525	6	E07164
33	486.8	44.5	525	6	E15131
34	486.8	44.5	525	6	I71150
35	486.8	44.5	525	6	I83709
36	486.8	44.5	525	22	E09431
37	486.8	44.5	660	9	H0MGSF
38	486.8	44.5	1498	9	HSGCSFRI
39	486.8	44.5	1521	6	E01219
40	486.8	44.5	1525	6	I05333
41	486.8	44.5	1525	6	I08669
42	486.8	44.5	1525	6	I09205
43	485.2	44.3	522	6	E01631
44	485.2	44.3	522	6	E08529
45	485.2	44.3	1414	9	H0MGSFB

ALIGNMENTS

RESULT	1				
E02574					
LOCUS	E02574	525 bp	DNA	PAT	29-SEP-1997
DEFINITION	DNA encoding human colony-stimulating factor derivative.				
ACCESSION	E02574				
VERSION	E02574.1	GI:2170804			
KEYWORDS	JP 1990227075-A/2.				
SOURCE	synthetic construct				
ORGANISM	artificial sequence.				
REFERENCE	1 (bases 1 to 525)				
AUTHORS	Sasaki, K., Nishi, T., Yasumura, S., Sato, M. and Itou, S.				
TITLE	NEW POLYPEPTIDE				
JOURNAL	Patent: JP 1990227075-A 2 10-SEP-1990;				
COMMENT	KYOWA HAKKO KOGYO CO LTD				
	OS Artificial gene				
	OC Artificial sequence; Genes.				
	PN JP 1990227075-A/2				
	PD 10-SEP-1990				
	PF 28-SEP-1989 JP 1989253097				

```
PR 29-SEP-1988 JP 88P 245705
PI SASAKI KATSUTOSHI, NISHI TATSUYA, YASUMURA SHIGEYOSHI, PI
SATO MORIYUKI,
PI ITOU SEIGA
PC C12N9/72,C07K13/00,C07K15/14,C12N1/21,C12N5/10,C12N9/64, PC
C12N15/27,
PC C12N15/58,C12N15/70,C12N15/85,C12P21/02,C12P21/02,(C12N1/21,
C12R1:19),
PC (C12P21/02,C12R1:19),(C12P21/02,C12R1:91);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
FH Key Location/Qualifiers
FH mat_peptide 1..522
FT /product='human colony-stimulating factor' FT
FT derivative which
FT is named 'hg-CSF[ND28]'.
FEATURES
source
Location/Qualifiers
1..525
/organism='synthetic construct'
/db_xref='taxon:32630' 99 t
BASE COUNT 94 a 184 c 148 g
ORIGIN
Query Match 47.5%; Score 520.4; DB 6; Length 525;
Best Local Similarity 99.8%; Pred. No. 3.3e-96;
Matches 521; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 574 GCGCCACATATCGCGCTCGAGTCTACACAGAGCTTCTTTTAAAGCTTAGAGCAA 633
Db 1 GCACCAACATATCGCGCTCGAGTCTACACAGAGCTTCTTTTAAAGCTTAGAGCAA 60
Qy 634 GTGAGGAGATCCAGGCGATGGCGAGCGCTCCAGAGAGAGCTGTGTGCCACTACAAG 693
Db 61 GTGAGGAGATCCAGGCGATGGCGAGCGCTCCAGAGAGAGCTGTGTGCCACTACAAG 120
Qy 694 CTGTGCCACCCCGAGGAGCTGTGTCTCGGACACTCTCTGGGCATCCCTGGGCTCCC 753
Db 121 CTGTGCCACCCCGAGGAGCTGTGTCTCGGACACTCTCTGGGCATCCCTGGGCTCCC 180
Qy 754 CTGAGCAGCTGCCCGACCCAGCGCCCTGCAGCTGGCAGCTGTGTGAGCCAACTCCATAGC 813
Db 181 CTGAGCAGCTGCCCGACCCAGCGCCCTGCAGCTGGCAGCTGTGTGAGCCAACTCCATAGC 240
Qy 814 GGCCTTTTCTCTACCAAGGGCTCTGTGAGGCGCTTGAAGGGATCTCCCGCGAGTTGGGT 873
Db 241 GGCCTTTTCTCTACCAAGGGCTCTGTGAGGCGCTTGAAGGGATCTCCCGCGAGTTGGGT 300
Qy 874 CCCACCTTGGACACACTGCAGCTGGAGTGGCGGACTTTGCCACACCATCTGGCAGCAG 933
Db 301 CCCACCTTGGACACACTGCAGCTGGAGTGGCGGACTTTGCCACACCATCTGGCAGCAG 360
Qy 934 ATGAAGAAGCTGGAATGGCCCTTGCCTGCAGCCACCCAGGGTGCCTATGCCGGCTTTC 993
Db 361 ATGAAGAAGCTGGAATGGCCCTTGCCTGCAGCCACCCAGGGTGCCTATGCCGGCTTTC 420
Qy 994 GCCTCTCTTCCAGCGCGGGGAGAGGGGTCTAGTTCCTTCCATCTGCAGAGCTTTC 1053
Db 421 GCCTCTCTTCCAGCGCGGGGAGAGGGGTCTAGTTCCTTCCATCTGCAGAGCTTTC 480
Qy 1054 CTGAGGAGTGTGTAACCGCTTCTACGCGCACTTGGCCAGGCC 1095
Db 481 CTGAGGAGTGTGTAACCGCTTCTACGCGCACTTGGCCAGGCC 522
RESULT 2
ID E11965
XX standard; RNA; HUM; 555 BP.
AC E11965;
```

```
XX E11965.1
SV 08-OCT-1997 (Rel. 52, Created)
DT 02-SEP-2000 (Rel. 65, Last updated, Version 2)
XX Human cDNA encoding thrombopoietin.
XX JP 1996228781-A/6.
XX Homo sapiens (human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
XX [1]
RP 1-555
RA Miyazaki H., Kato H., Ogami K., Iwamatsu A., Akahori H., Kuroki R.,
RA Shimizu T., Muto T.;
RT "DNA CODING FOR PROTEIN HAVING TPO ACTIVITY";
RL Patent number JP1996228781-A/6, 10-SEP-1996.
RL KIRIN BREWERY CO LTD.
XX OS Homo sapiens (human)
CC PN JP 1996228781-A/6
CC PD 10-SEP-1996
CC PF 14-FEB-1995 JP 1995063298
CC PR 14-FEB-1994 JP 94P 39090, 25-MAR-1994 JP 94P 79842,
CC PR 01-JUN-1994 JP 94P 155126, 15-JUN-1994 JP 94P 167328,
CC PR 17-AUG-1994 JP 94P 227159, 17-AUG-1994 JP 94P 193169,
CC PR 01-NOV-1994 JP 94P 304167, 01-DEC-1994 JP 94P 298669,
CC PR 28-DEC-1994 JP 94P 341200
CC PI MIYAZAKI HIROSHI, KATO HISASHI, OGAMI KINYA, IWAMATSU AKIHIKO,
CC PI AKAHORI HIROHORI, KUROKI RYOTA, SHIMIZU TOSHIYUKI,
CC PI MUTO TAKANORI
CC PC C12N15/09,A61K38/00,C07K14/52,C12N1/21,C12N5/10,C12P21/02,
CC PC (C12N1/21,
CC PC C12R1:19),(C12N5/10,C12R1:91),(C12P21/02,C12R1:19),(C12P21/02,
CC PC C12R1:91);
CC CC strandedness: Double;
CC CC topology: Linear;
CC FH Key Location/Qualifiers
CC FH source 1..555
CC FT /organism='Homo sapiens'
CC FT /tissue_type='liver'
CC FT CDS 1..555
CC FT /product='thrombopoietin'
CC FT sig_peptide 1..63
XX Key Location/Qualifiers
FH source 1..555
FH /db_xref='taxon:9606'
FH /organism='Homo sapiens'
SQ Sequence 555 BP; 106 A; 183 C; 145 G; 121 T; 0 other;
Query Match 47.4%; Score 518.8; DB 22; Length 555;
Best Local Similarity 99.6%; Pred. No. 7e-96;
Matches 520; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 ATGGAGCTGACTGAATTGCTCGTGGTTCATGCTCTCTCTAACTCAAGGCTAACGCTG 60
Db 1 ATGGAGCTGACTGAATTGCTCGTGGTTCATGCTCTCTCTAACTCAAGGCTAACGCTG 60
Qy 61 TCCAGCCCGGCTCCTCTCTGTTGTGACCTCCGAGTCCCTAGTAACTGCTTCGTGACTCC 120
Db 61 TCCAGCCCGGCTCCTCTCTGTTGTGACCTCCGAGTCCCTAGTAACTGCTTCGTGACTCC 120
Qy 121 CATGTCTTTCACAGCAGACTGAGCCAGCTGCCAGAGGTTCCACCTTTGCGCTACACCTGTC 180
Db 121 CATGTCTTTCACAGCAGACTGAGCCAGGTTCCACCTTTGCGCTACACCTGTC 180
```

QY 181 CTGCTGCTGTGTGGAGCTTTAGCTTGGAGAAATGGAAACCCAGATGGAGGACCAAG 240
|||||
Db 181 CTGCTGCTGTGTGGAGCTTTAGCTTGGAGAAATGGAAACCCAGATGGAGGACCAAG 240
QY 241 GCACAGACATTTCTGGAGAGTGAACCTTCTGTCTGGAGGAGTGTGGAGCAGCGGGA 300
|||||
Db 241 GCACAGACATTTCTGGAGAGTGAACCTTCTGTCTGGAGGAGTGTGGAGCAGCGGGA 300
QY 301 CAACTGGGACCACTTCCCTCTCATCCCTCTCTGGGGCAGCTTCTTGACAGGTCCTCTC 360
|||||
Db 301 CAACTGGGACCACTTCCCTCTCATCCCTCTCTGGGGCAGCTTCTTGACAGGTCCTCTC 360
QY 361 CTCTTGGGGCCCTGCAGAGCCTCTTGGAAACCAGCTTCTTCCACAGGGCAGGACCA 420
Db 361 CTCTTGGGGCCCTGCAGAGCCTCTTGGAAACCAGCTTCTTCCACAGGGCAGGACCA 420
QY 421 GCTCACAAGATCCCAATGCCATCTTCTGAGCTTCCAAACCTGTCCGAGGAAAGGTG 480
|||||
Db 421 GCTCACAAGATCCCAATGCCATCTTCTGAGCTTCCAAACCTGTCCGAGGAAAGGTG 480
QY 481 CGTTTCTGTAGCTTGTAGGAGGTCACACCTCTGCGTACGG 522
|||||
Db 481 CGTTTCTGTAGCTTGTAGGAGGTCACACCTCTGCGTACGG 522

RESULT 3
AR103959
LOCUS AR103959 1059 bp DNA PAT 14-FEB-2001
DEFINITION Sequence 57 from patent US 6093395.
ACCESSION AR103959
VERSION AR103959.1 GI:12816667
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1059)
AUTHORS Bauer, S., Christopher, Abrams, M., Allen, Bradford-Goldberg, S., Ruth, Caparon, M., Helena, Easton, A., Michael, Klein, B., Kure, McKearn, J. P., Ollins, P. O., Paik, K. and Thomas, J. W.
TITLE Co-administration of interleukin-3 mutant polypeptides with CSF's for multi-lineage hematopoietic cell production
JOURNAL Patent: US 6093395-A 57 25-JUL-2000;
FEATURES Location/Qualifiers
source 1..1059
BASE COUNT 221 a 375 c 236 g 227 t
ORIGIN

Query Match 47.4%; Score 518.8; DB 6; Length 1059;
Best Local Similarity 99.6%; Pred. No. 5.8e-96;
Matches 520; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGAGCTGACTGAATTCCTCGTGGTTCATGCTTCTCCTAACTGCAAGGCTAACGCTG 60
|||||
Db 1 ATGGAGCTGACTGAATTCCTCGTGGTTCATGCTTCTCCTAACTGCAAGGCTAACGCTG 60
QY 61 TCAGCCCGGCTCTCTGTGTGTGACCTCCGAGTCTCAGTAACTGCTTCGTGACTCC 120
|||||
Db 61 TCAGCCCGGCTCTCTGTGTGTGACCTCCGAGTCTCAGTAACTGCTTCGTGACTCC 120
QY 121 CATGTCTTCCACAGACAGTACGAGTCCAGAGTTCACACCTTTGGCTACACCTGTC 180
|||||
Db 121 CATGTCTTCCACAGACAGTACGAGTCCAGAGTTCACACCTTTGGCTACACCTGTC 180
QY 181 CTGCTGCTGCTGTGGACTTTAGCTTGGAGAAATGGAAACCCAGATGGAGGACCAAG 240
|||||
Db 181 CTGCTGCTGCTGTGGACTTTAGCTTGGAGAAATGGAAACCCAGATGGAGGACCAAG 240
QY 241 GCACAGACATTTCTGGAGAGTGAACCTTCTGTCTGGAGGAGTGTGGAGCAGCGGGA 300
|||||
Db 241 GCACAGACATTTCTGGAGAGTGAACCTTCTGTCTGGAGGAGTGTGGAGCAGCGGGA 300

QY 301 CAACTGGGACCACTTCCCTCTCATCCCTCTCTGGGGCAGCTTCTTGACAGGTCCTCTC 360
|||||
Db 301 CAACTGGGACCACTTCCCTCTCATCCCTCTCTGGGGCAGCTTCTTGACAGGTCCTCTC 360
QY 361 CTCTTGGGGCCCTGCAGAGCCTCTTGGAAACCAGCTTCTTCCACAGGGCAGGACCA 420
Db 361 CTCTTGGGGCCCTGCAGAGCCTCTTGGAAACCAGCTTCTTCCACAGGGCAGGACCA 420
QY 421 GCTCACAAGATCCCAATGCCATCTTCTGAGCTTCCAAACCTGTCCGAGGAAAGGTG 480
|||||
Db 421 GCTCACAAGATCCCAATGCCATCTTCTGAGCTTCCAAACCTGTCCGAGGAAAGGTG 480
QY 481 CGTTTCTGTAGCTTGTAGGAGGTCACACCTCTGCGTACGG 522
|||||
Db 481 CGTTTCTGTAGCTTGTAGGAGGTCACACCTCTGCGTACGG 522

RESULT 4
AR087133
LOCUS AR087133 1062 bp DNA PAT 07-SEP-2000
DEFINITION Sequence 3 from patent US 5986049.
ACCESSION AR087133
VERSION AR087133.1 GI:10013899
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1062)
AUTHORS Forstrom, J. W., Lofton-Day, C. E. and Lok, S.
TITLE Purified thrombopoietin and method of making it
JOURNAL Patent: US 5986049-A 3 16-NOV-1999;
FEATURES Location/Qualifiers
source 1..1062
BASE COUNT 223 a 375 c 236 g 228 t
ORIGIN

Query Match 47.4%; Score 518.8; DB 6; Length 1062;
Best Local Similarity 99.6%; Pred. No. 5.8e-96;
Matches 520; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGAGCTGACTGAATTCCTCGTGGTTCATGCTTCTCCTAACTGCAAGGCTAACGCTG 60
|||||
Db 1 ATGGAGCTGACTGAATTCCTCGTGGTTCATGCTTCTCCTAACTGCAAGGCTAACGCTG 60
QY 61 TCAGCCCGGCTCTCTGTGTGTGACCTCCGAGTCTCAGTAACTGCTTCGTGACTCC 120
|||||
Db 61 TCAGCCCGGCTCTCTGTGTGTGACCTCCGAGTCTCAGTAACTGCTTCGTGACTCC 120
QY 121 CATGTCTTCCACAGACAGTACGAGTCCAGAGTTCACACCTTTGGCTACACCTGTC 180
|||||
Db 121 CATGTCTTCCACAGACAGTACGAGTCCAGAGTTCACACCTTTGGCTACACCTGTC 180
QY 181 CTGCTGCTGCTGTGGACTTTAGCTTGGAGAAATGGAAACCCAGATGGAGGACCAAG 240
|||||
Db 181 CTGCTGCTGCTGTGGACTTTAGCTTGGAGAAATGGAAACCCAGATGGAGGACCAAG 240
QY 241 GCACAGACATTTCTGGAGAGTGAACCTTCTGTCTGGAGGAGTGTGGAGCAGCGGGA 300
|||||
Db 241 GCACAGACATTTCTGGAGAGTGAACCTTCTGTCTGGAGGAGTGTGGAGCAGCGGGA 300
QY 301 CAACTGGGACCACTTCCCTCTCATCCCTCTCTGGGGCAGCTTCTTGACAGGTCCTCTC 360
|||||
Db 301 CAACTGGGACCACTTCCCTCTCATCCCTCTCTGGGGCAGCTTCTTGACAGGTCCTCTC 360
QY 361 CTCTTGGGGCCCTGCAGAGCCTCTTGGAAACCAGCTTCTTCCACAGGGCAGGACCA 420
Db 361 CTCTTGGGGCCCTGCAGAGCCTCTTGGAAACCAGCTTCTTCCACAGGGCAGGACCA 420
QY 421 GCTCACAAGATCCCAATGCCATCTTCTGAGCTTCCAAACCTGTCCGAGGAAAGGTG 480
|||||

Db 421 GCTCACAGGATCCCAATGCATCTTCTCGAGCTTCCACACACCTGCTCCGAGGAAAGGTG 480
Qy 481 CGTTTCTGATGCTGTAGGAGGTCACACCTCTCGGTAGG 522
Db 481 CGTTTCTGATGCTGTAGGAGGTCACACCTCTCGGTAGG 522

RESULT 5

AR087571 LOCUS AR087571 1062 bp DNA PAT 07-SEP-2000
DEFINITION Sequence 18 from patent US 5989537.
ACCESSION AR087571
VERSION AR087571.1 GI:10014334
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.

REFERENCE 1 (bases 1 to 1062)
AUTHORS Holly,R.D., Lok,S., Foster,D.C., Hagen,F.S., Kaushansky,K.,
Kuijper,J.L., Lofton-Day,C.E. and Oort,P.J.
TITLE Methods for stimulating granulocyte/macrophage lineage using
thrombopoietin
JOURNAL Patent: US 5989537-A 18 23-NOV-1999;
FEATURES Location/Qualifiers
Source 1..1062
/organism="unknown"

BASE COUNT 223 a 375 c 236 g 228 t
ORIGIN

Query Match 47.4%; Score 518.8; DB 6; Length 1062;
Best Local Similarity 99.6%; Pred. No. 5.8e-96;
Matches 520; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGGAGCTGACTGAATGCTCTCGTGGTCATGCTTCTCTAACTCAAGGCTAAGCGTG 60
Db 1 ATGGAGCTGACTGAATGCTCTCGTGGTCATGCTTCTCTAACTCAAGGCTAAGCGTG 60
Qy 61 TCCAGCCCGGCTCCTCTGCTGTGACCTCCGAGTCCCTCAGTAAACTGCTTCGTGACTCC 120
Db 61 TCCAGCCCGGCTCCTCTGCTGTGACCTCCGAGTCCCTCAGTAAACTGCTTCGTGACTCC 120
Qy 121 CATGCTCTCACAGCAGACTGAGCCAGTCCGAGAGTTTCAACCTTTGCTTACACTGTC 180
Db 121 CATGCTCTCACAGCAGACTGAGCCAGTCCGAGAGTTTCAACCTTTGCTTACACTGTC 180
Qy 181 CTGCTGCTCTGTGACCTTTAGCTTGGGAGATGGAAACCCAGATGGAGGACCAAG 240
Db 181 CTGCTGCTCTGTGACCTTTAGCTTGGGAGATGGAAACCCAGATGGAGGACCAAG 240
Qy 241 GCACAGACATCTGGGAGCAGTACCTTCTGCTGGAGGAGTATGGCAGCAGGGGA 300
Db 241 GCACAGACATCTGGGAGCAGTACCTTCTGCTGGAGGAGTATGGCAGCAGGGGA 300
Qy 301 CAACTGGGACCCACTTGCCTTCATCCCTCTGCTGGGCGACCTTTCTGGACAGTCCGTC 360
Db 301 CAACTGGGACCCACTTGCCTTCATCCCTCTGCTGGGCGACCTTTCTGGACAGTCCGTC 360
Qy 361 CTCTCTGGGCGCTCGAGAGCTTCTTGAACCCAGCTTCTTCCACAGGAGGACCAACA 420
Db 361 CTCTCTGGGCGCTCGAGAGCTTCTTGAACCCAGCTTCTTCCACAGGAGGACCAACA 420
Qy 421 GCTCACAAGATCCCAATGCCATCTTCTGAGCTTCCACACCTGCTCCGAGGAAAGGTG 480
Db 421 GCTCACAAGATCCCAATGCCATCTTCTGAGCTTCCACACCTGCTCCGAGGAAAGGTG 480
Qy 481 CGTTTCTGATGCTGTAGAGGGTCCACCTCTCGGTACGG 522
Db 481 CGTTTCTGATGCTGTAGAGGGTCCACCTCTCGGTACGG 522

RESULT 6

I49760

LOCUS I49760 1062 bp DNA PAT 07-OCT-1997
DEFINITION Sequence 3 from patent US 5641655.
ACCESSION I49760
VERSION I49760.1 GI:2471980
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.

REFERENCE 1 (bases 1 to 1062)
AUTHORS Foster,D.C., Helpel,M.D. and Holly,R.D.
TITLE Methods for producing thrombopoietin polypeptides using a mammalian
tissue plasminogen activator secretory peptide
JOURNAL Patent: US 5641655-A 3 24-JUN-1997;
FEATURES Location/Qualifiers
Source 1..1062
/organism="unknown"

BASE COUNT 223 a 375 c 236 g 228 t
ORIGIN

Query Match 47.4%; Score 518.8; DB 6; Length 1062;
Best Local Similarity 99.6%; Pred. No. 5.8e-96;
Matches 520; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGGAGCTGACTGAATGCTCTCGTGGTCATGCTTCTCTAACTCAAGGCTAAGCGTG 60
Db 1 ATGGAGCTGACTGAATGCTCTCGTGGTCATGCTTCTCTAACTCAAGGCTAAGCGTG 60
Qy 61 TCCAGCCCGGCTCCTCTGCTGTGACCTCCGAGTCCCTCAGTAAACTGCTTCGTGACTCC 120
Db 61 TCCAGCCCGGCTCCTCTGCTGTGACCTCCGAGTCCCTCAGTAAACTGCTTCGTGACTCC 120
Qy 121 CATGCTCTCACAGCAGACTGAGCCAGTCCGAGAGTTTCAACCTTTGCTTACACTGTC 180
Db 121 CATGCTCTCACAGCAGACTGAGCCAGTCCGAGAGTTTCAACCTTTGCTTACACTGTC 180
Qy 181 CTGCTGCTCTGTGACCTTTAGCTTGGGAGATGGAAACCCAGATGGAGGACCAAG 240
Db 181 CTGCTGCTCTGTGACCTTTAGCTTGGGAGATGGAAACCCAGATGGAGGACCAAG 240
Qy 241 GCACAGGACATCTGGGAGCAGTACCTTCTGCTGGAGGAGTATGGCAGCAGGGGA 300
Db 241 GCACAGGACATCTGGGAGCAGTACCTTCTGCTGGAGGAGTATGGCAGCAGGGGA 300
Qy 301 CAACTGGGACCCACTTGCCTTCATCCCTCTGCTGGGCGACCTTTCTGGACAGTCCGTC 360
Db 301 CAACTGGGACCCACTTGCCTTCATCCCTCTGCTGGGCGACCTTTCTGGACAGTCCGTC 360
Qy 361 CTCTCTGGGCGCTCGAGAGCTTCTTGAACCCAGCTTCTTCCACAGGAGGACCAACA 420
Db 361 CTCTCTGGGCGCTCGAGAGCTTCTTGAACCCAGCTTCTTCCACAGGAGGACCAACA 420
Qy 421 GCTCACAAGATCCCAATGCCATCTTCTGAGCTTCCACACCTGCTCCGAGGAAAGGTG 480
Db 421 GCTCACAAGATCCCAATGCCATCTTCTGAGCTTCCACACCTGCTCCGAGGAAAGGTG 480
Qy 481 CGTTTCTGATGCTGTAGAGGGTCCACCTCTCGGTACGG 522
Db 481 CGTTTCTGATGCTGTAGAGGGTCCACCTCTCGGTACGG 522

RESULT 7

HUMTHROMB LOCUS HUMTHROMB 1062 bp mRNA PRI 22-DEC-1994
DEFINITION Human thrombopoietin mRNA, complete cds.
ACCESSION L36052
VERSION L36052.1 GI:533216
KEYWORDS thrombopoietin.
SOURCE Homo sapiens cdna to mRNA.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1062)
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

ACCESSION	EL6669
VERSION	EL6669.1
KEYWORDS	GI:5711352
SOURCE	JP 1998212243-A/3.
ORGANISM	Homo sapiens.
	Homo sapiens
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE	1 (bases 1 to 1086)
AUTHORS	Kurda, K
TITLE	COMPOSITION CONTAINING THROMBOPOIETIN PROTEIN AND USED FOR STIMULATING RECOVERY OF HEMOSTATIC THROMBUS-FORMING ABILITY
JOURNAL	Patent: JP 1998212243-A 3 11-AUG-1998;
COMMENT	KIRIN BREWERY CO LTD OS Homo sapiens (human) PN JP 1998212243-A/3 PD 11-AUG-1998

31-JAN-1997 JP 1997019549
KURODA KENJI
A61K38/00, A01N1/02, C07K14/52//C12N15/09, C12P21/02, C12P21/02,
C12R1/19;
strandedness: Double;
topology: Linear;
hypothetical: No;
anti-sense: No;

Key	Location/Qualifiers
FH	
FH	
FT	1..1086
FT	/organism='Homo sapiens'
FT	/tissue_type='Liver'
FT	1..24
FT	5'UTR
FT	25..87
FT	sig_peptide

FT	mat_peptide	88..1086	/product='signal peptide of thrombopoietin'
FT	CDS	25..1086	/product='thrombopoietin'
FT			/product='thrombopoietin'

```
FEATURES
  Location/Qualifiers
    1..1086
      /organism="Homo sapiens"
```

```

BASE COUNT      229 a 386 c 243 g 228 t
ORIGIN
/db_xref="taxon:9606"

```

Query Match	47.4%	Score	518	8	DB	6	Length	1086
-------------	-------	-------	-----	---	----	---	--------	------

Best Local Similarity	99.68;	Pred. No. 5.8e-96;		
Matches 520;	Conservative	0;	Mismatches 2;	Indels 0;
Gaps	0;			

QY 1 ATGGAGCTGACTGAATTGCTTCTCCTGGTCATGCTTCTCTAACTGCAAGCGTAAACGCTG 60

D0	25	ATGAGAGCTGACTGAAATGCTCTCTGTTGGTGATGCTTCTTCTTAACATGCAAGGCTATACGCTG	84
QY	61	TCCAGCCCCGCTCTCTCTGTTGTGACCTCCGAGTCCCTCAGTAACCTGCTTCGTGACTCC	120

Db 85 TCCAGCCGGCTCCTCTCTGCTTGTGACCTCCGAGTCCTCAGTAACTGCTTCGTGACTCC 144

	QY	121 CATGTCCCTTACAGCAGACTGAGCGAGTGCCCCAGAGGGTTACCCTTGGCTACACCTGTC	180
D6		145 CATGTCCCTTACAGCAGACTGAGCGAGTGCCCCAGAGGGTTACCCTTGGCTACACCTGTC	204

QY 181 CTGCTGCCTGCTGTGGACTTTAGCTTGGAGAAATGGAAACCCAGATGGAGAGACCAG 240
|||

Db 205 CTGCTGCCCTGCTGTGGACATTAGCTTGGGAGATCGAAACCCAGATGGAGGAGACCAAG 264

QY 241 GCACAGGACATTCGGGAGCAGTGACCCCTTCTGCTGGAGGAGTGTATGCCAGCACGGGA 300

Db 265 GCACAGACATTTGGGAGCAGTGACCCCTTCTGCTGGAGGAGTGATGGCAGCACGGGA 324

QY	301	CAACTGGGACCCACTTGCCTCTCATCCCTCTGGGCGACGTTCTGGACAGGTCCGCTC	360
Dh	325	CAATGGGACCCACTTGCCTCTCATCCCTCTGGGCGACGTTCTGGACAGGTCCGCTC	384

361 CTCCTTGGGGCCCTGCAGAGCCTCCTTGGAAACCCAGCTTCCTCCACAGGCAGACCACA 420

```

|||||
385 CTCCTGGGCGCTGCAGAGCCTCTTGGAAACCCAGCTCTCCACAGGCGAGGCCACCA 444
|||||
421 GCTCACAGAGTCCCATGCTTCCTTCAGCTTCCACACACCTGCTCCGAGGAAGGTG 480
|||||
445 GCTCACAGAGTCCCATGCTTCCTTCAGCTTCCACACACCTGCTCCGAGGAAGGTG 504
|||||
481 CGTTTCCTGATGCTTGTAGGAGGCTCCACCTCTCGGTACGG 522
|||||
505 CGTTTCCTGATGCTTGTAGGAGGCTCCACCTCTCGGTACGG 546
|||||

RESULT 9
LOCUS E12182 1267 bp DNA PAT 24-JUN-1998
DEFINITION Human cDNA encoding thrombopoietin.
ACCESSION E12182.1 GI:3251016
VERSION JP 1996277296-A/2.
KEYWORDS Homo sapiens.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1267)
Miyazaki,H., Kato,H., Ogami,K., Iwamatsu,A., Akahori,H., Kuroki,R.,
Shimizu,T. and Muto,T.
PROTEIN HAVING TPO ACTIVITY
Patent: JP 1996277296-A 2 22-OCT-1996;
JOURNAL KIRIN BREWERY CO LTD
COMMENT OS Homo sapiens (human)
PN JP 1996277296-A/2
PD 22-OCT-1996
PF 14-FEB-1995 JP 1995161363
PR 14-FEB-1994 JP 94P 39090, 25-MAR-1994 JP 94P 79842, PR
01-JUN-1994 JP 94P 155126, 15-JUN-1994 JP 94P 167328, PR
17-AUG-1994 JP 94P 227159, 17-AUG-1994 JP 94P 193169, PR
01-NOV-1994 JP 94P 304167, 01-DEC-1994 JP 94P 298669, PR
28-DEC-1994 JP 94P 341200
PI MIYAZAKI HIROSHI, KATO HISASHI, OGAMI KINYA, IWAMATSU AKIHIKO,
PI AKAHORI HIRONORI, KUROKI RYOTA, SHIMIZU TOSHIYUKI, PI MUTO
TAKANORI
PC C07K14/52,C07H21/04,C12N15/09,C12P21/02//A61K38/00,(C12P21/02,
C12R1:19),
PC (C12P21/02,C12R1:91);
CC strandedness: Double;
CC topology: Linear;
FH key Location/Qualifiers
FH source 1..1267
FT /organism='Homo sapiens'
FT /tissue_type='liver'
FT CDS 25..1086
FT /product='thrombopoietin'
FT sig_peptide 25..87
FEATURES
source 1..1267
/organism='Homo sapiens'
/db_xref='taxon:9606'
BASE COUNT 281 a 436 c 279 g 271 t
ORIGIN

Query Match 47.4%; Score 518.8; DB 6; Length 1267;
Best Local Similarity 99.6%; Pred. No. 5.6e-96;
Matches 520; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGGAGCTGACTGAATGCTTCCTCGTGGTCATGCTTCTCCTAACTGCAAGGCTAACGCTG 60
|||||
Db 25 ATGGAGCTGACTGAATGCTTCCTCGTGGTCATGCTTCTCCTAACTGCAAGGCTAACGCTG 84
|||||
Qy 61 TCCAGCCGGCTCCTCGTGGTGTGACCTCCGAGTCCCTCAGTAAACTGCTTCGTGACTCC 120
|||||
Db 85 TCCAGCCGGCTCCTCGTGGTGTGACCTCCGAGTCCCTCAGTAAACTGCTTCGTGACTCC 144
|||||

```

```

Qy 121 CATGCTCTTACAGCAGACTGAGCCAGTGCACAGAGGTTCCACCTTTGCCTACACCTGTC 180
|||||
Db 145 CATGCTCTTACAGCAGACTGAGCCAGTGCACAGAGGTTCCACCTTTGCCTACACCTGTC 204
|||||
Qy 181 CTGCTGCCCTGCTGTGGACATTTAGCTTTGGAGAAATGAAACCCAGATGGAGGAGACCAAG 240
|||||
Db 205 CTGCTGCCCTGCTGTGGACATTTAGCTTTGGAGAAATGAAACCCAGATGGAGGAGACCAAG 264
|||||
Qy 241 GCACAGGACATTTCTGGAGCAGTGCACCTTCTGCTGGAGGAGTGTGGCAGCAGCGGGA 300
|||||
Db 265 GCACAGGACATTTCTGGAGCAGTGCACCTTCTGCTGGAGGAGTGTGGCAGCAGCGGGA 324
|||||
Qy 301 CAACTGGGACCCACATTCCTCTCATCCCTCTGGGCGAGCTTTCTGTGACAGGTCCTGCTC 360
|||||
Db 325 CAACTGGGACCCACATTCCTCTCATCCCTCTGGGCGAGCTTTCTGTGACAGGTCCTGCTC 384
|||||
Qy 361 CTCTTTGGGGCCCTGACAGAGCTCTCTTTGGAACCCAGCTTCCTCCACAGGCGAGACCAACA 420
|||||
Db 385 CTCTTTGGGGCCCTGACAGAGCTCTCTTTGGAACCCAGCTTCCTCCACAGGCGAGACCAACA 444
|||||
Qy 421 GCTCAAGGATCCCAATGCCATCTTCTGAGCTTCCACACCTGCTCCGAGGAAAGGTG 480
|||||
Db 445 GCTCAAGGATCCCAATGCCATCTTCTGAGCTTCCACACCTGCTCCGAGGAAAGGTG 504
|||||
Qy 481 CGTTTCCTGATGCTTGTAGGAGGTCACCCCTCTCGGTACGG 522
|||||
Db 505 CGTTTCCTGATGCTTGTAGGAGGTCACCCCTCTCGGTACGG 546
|||||

RESULT 10
LOCUS E12214 1267 bp DNA PAT 24-JUN-1998
DEFINITION Human cDNA encoding thrombopoietin.
ACCESSION E12214
VERSION E12214.1 GI:3251048
KEYWORDS JP 1996291196-A/1.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1267)
Miyazaki,H., Kato,H., Ogami,K., Iwamatsu,A., Akahori,H., Kuroki,R.,
Shimizu,T. and Muto,T.
PROTEIN HAVING TPO ACTIVITY
Patent: JP 1996291196-A 1 05-NOV-1996;
JOURNAL KIRIN BREWERY CO LTD
COMMENT OS Homo sapiens (human)
PN JP 1996291196-A/1
PD 05-NOV-1996
PF 14-FEB-1995 JP 1995355052
PR 14-FEB-1994 JP 94P 39090, 25-MAR-1994 JP 94P 79842, PR
01-JUN-1994 JP 94P 155126, 15-JUN-1994 JP 94P 167328, PR
17-AUG-1994 JP 94P 227159, 17-AUG-1994 JP 94P 193169, PR
01-NOV-1994 JP 94P 304167, 01-DEC-1994 JP 94P 298669, PR
28-DEC-1994 JP 94P 341200
PI MIYAZAKI HIROSHI, KATO HISASHI, OGAMI KINYA, IWAMATSU AKIHIKO,
PI AKAHORI HIRONORI, KUROKI RYOTA, SHIMIZU TOSHIYUKI, PI MUTO
TAKANORI
PC C07K14/53,C07K1/22,C12N15/09//A61K38/00,C12P21/08; CC
strandedness: Double;
CC topology: Linear;
FH key Location/Qualifiers
FH source 1..1267
FT /organism='Homo sapiens'
FT /tissue_type='liver'
FT CDS 25..1086
FT /product='thrombopoietin'
FT sig_peptide 25..87
FT mat_peptide 88..1083
FEATURES
source 1..1267
/organism='Homo sapiens'
/db_xref='taxon:9606'

```

```
source 1..1267
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 281 a 436 c 279 g 271 t
ORIGIN
Query Match 47.4%; Score 518.8; DB 6; Length 1267;
Best Local Similarity 99.6%; Pred. No. 5.5e-96;
Matches 520; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAGCTGACTGAATGCTCCGTCGTCATGCTTCTCCTAACTGAAGCTAACGCTG 60
Db ATGAGCTGACTGAATGCTCCGTCGTCATGCTTCTCCTAACTGAAGCTAACGCTG 95

QY 61 TCAGAGCCGGCTCCTCTGCTGTTGACCTCCGAGTCCAGTAACTGCTTCGTGACTCC 120
Db TCAGAGCCGGCTCCTCTGCTGTTGACCTCCGAGTCCAGTAACTGCTTCGTGACTCC 155

QY 121 CATGTCCTTCCACAGCAGACTGAGCCAGTGCAGAGTTCACCTTTGCTTACACCTGTC 180
Db CATGTCCTTCCACAGCAGACTGAGCCAGTGCAGAGTTCACCTTTGCTTACACCTGTC 215

QY 181 CTGCTGCTCTGCTGGAGTCTTAGCTTGGGAGATGGAACCCAGATGGAGGACCAAG 240
Db CTGCTGCTCTGCTGGAGTCTTAGCTTGGGAGATGGAACCCAGATGGAGGACCAAG 275

QY 241 GCACAGACATTCCTGGAGCAGTGACCTTCTGCTGGAGGAGTGATGGCAGCACGGGA 300
Db GCACAGACATTCCTGGAGCAGTGACCTTCTGCTGGAGGAGTGATGGCAGCACGGGA 335

QY 301 CAACTGGGACCCACTTGCCTCTCATCCCTCTGGGCGAGCTTCTGGACAGGTCCTGTC 360
Db CAACTGGGACCCACTTGCCTCTCATCCCTCTGGGCGAGCTTCTGGACAGGTCCTGTC 395

QY 361 CTCTTTGGGCGCTCGCAGAGCTCTTGGAAACCCAGCTTCTCCACAGGCGAGGACACA 420
Db CTCTTTGGGCGCTCGCAGAGCTCTTGGAAACCCAGCTTCTCCACAGGCGAGGACACA 455

QY 421 GCTCACAAGATCCCAATGCCATCTTCTGAGCTTCCACACCTGCTCCAGGAAAGGTG 480
Db GCTCACAAGATCCCAATGCCATCTTCTGAGCTTCCACACCTGCTCCAGGAAAGGTG 515

QY 481 CGTTTCTGATGCTGTAGGAGGTCACCTCTGCGTACGG 522
Db CGTTTCTGATGCTGTAGGAGGTCACCTCTGCGTACGG 557

RESULT 12
AR023468 AR023468 1342 bp DNA PAT 05-DEC-1998
LOCUS Sequence 24 from patent US 5795569.
DEFINITION AR023468
ACCESSION AR023468
VERSION AR023468.1 GI:3976762
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1342)
AUTHORS Bartley,T.D., Bogenberger,J.M., Bosselman,R.A., Hunt,P.,
Kinstler,O.B. and Samal,B.B.
TITLE Mono-pegylated proteins that stimulate megakaryocyte growth and
differentiation
JOURNAL Patent: US 5795569-A 24 18-AUG-1998;
FEATURES Location/Qualifiers
SOURCE 1..1342
/organism="unknown"
BASE COUNT 299 a 454 c 293 g 296 t
ORIGIN
Query Match 47.4%; Score 518.8; DB 6; Length 1342;
Best Local Similarity 99.6%; Pred. No. 5.5e-96;
Matches 520; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAGCTGACTGAATGCTCCGTCGTCATGCTTCTCCTAACTGAAGCTAACGCTG 60
Db ATGAGCTGACTGAATGCTCCGTCGTCATGCTTCTCCTAACTGAAGCTAACGCTG 95

QY 61 TCAGAGCCGGCTCCTCTGCTGTTGACCTCCGAGTCCAGTAACTGCTTCGTGACTCC 120
Db TCAGAGCCGGCTCCTCTGCTGTTGACCTCCGAGTCCAGTAACTGCTTCGTGACTCC 155

QY 121 CATGTCCTTCCACAGCAGACTGAGCCAGTGCAGAGTTCACCTTTGCTTACACCTGTC 180
Db CATGTCCTTCCACAGCAGACTGAGCCAGTGCAGAGTTCACCTTTGCTTACACCTGTC 215

QY 181 CTGCTGCTCTGCTGGAGTCTTAGCTTGGGAGATGGAACCCAGATGGAGGACCAAG 240
Db CTGCTGCTCTGCTGGAGTCTTAGCTTGGGAGATGGAACCCAGATGGAGGACCAAG 275

QY 241 GCACAGACATTCCTGGAGCAGTGACCTTCTGCTGGAGGAGTGATGGCAGCACGGGA 300
Db GCACAGACATTCCTGGAGCAGTGACCTTCTGCTGGAGGAGTGATGGCAGCACGGGA 335

QY 301 CAACTGGGACCCACTTGCCTCTCATCCCTCTGGGCGAGCTTCTGGACAGGTCCTGTC 360
Db CAACTGGGACCCACTTGCCTCTCATCCCTCTGGGCGAGCTTCTGGACAGGTCCTGTC 384

QY 361 CTCTTTGGGCGCTCGCAGAGCTCTTGGAAACCCAGCTTCTCCACAGGCGAGGACACA 420
Db CTCTTTGGGCGCTCGCAGAGCTCTTGGAAACCCAGCTTCTCCACAGGCGAGGACACA 444

QY 421 GCTCACAAGATCCCAATGCCATCTTCTGAGCTTCCACACCTGCTCCAGGAAAGGTG 480
Db GCTCACAAGATCCCAATGCCATCTTCTGAGCTTCCACACCTGCTCCAGGAAAGGTG 504

QY 481 CGTTTCTGATGCTGTAGGAGGTCACCTCTGCGTACGG 522
Db CGTTTCTGATGCTGTAGGAGGTCACCTCTGCGTACGG 546

RESULT 11
AR008878 AR008878 1342 bp DNA PAT 04-DEC-1998
LOCUS Sequence 1 from patent US 5756083.
DEFINITION AR008878
ACCESSION AR008878
VERSION AR008878.1 GI:3967683
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1342)
AUTHORS Elliott,S.G.
TITLE Mpl ligand analogs
JOURNAL Patent: US 5756083-A 1 26-MAY-1998;
FEATURES Location/Qualifiers
SOURCE 1..1342
/organism="unknown"
BASE COUNT 299 a 454 c 293 g 296 t
ORIGIN
Query Match 47.4%; Score 518.8; DB 6; Length 1342;
Best Local Similarity 99.6%; Pred. No. 5.5e-96;
Matches 520; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 1 ATGAGCTGACTGAATGCTCCGTCGTCATGCTTCTCCTAACTGAAGCTAACGCTG 60
Db ATGAGCTGACTGAATGCTCCGTCGTCATGCTTCTCCTAACTGAAGCTAACGCTG 95

QY 61 TCAGAGCCGGCTCCTCTGCTGTTGACCTCCGAGTCCAGTAACTGCTTCGTGACTCC 120
Db TCAGAGCCGGCTCCTCTGCTGTTGACCTCCGAGTCCAGTAACTGCTTCGTGACTCC 155

QY 121 CATGTCCTTCCACAGCAGACTGAGCCAGTGCAGAGTTCACCTTTGCTTACACCTGTC 180
Db CATGTCCTTCCACAGCAGACTGAGCCAGTGCAGAGTTCACCTTTGCTTACACCTGTC 215

QY 181 CTGCTGCTCTGCTGGAGTCTTAGCTTGGGAGATGGAACCCAGATGGAGGACCAAG 240
Db CTGCTGCTCTGCTGGAGTCTTAGCTTGGGAGATGGAACCCAGATGGAGGACCAAG 275

QY 241 GCACAGACATTCCTGGAGCAGTGACCTTCTGCTGGAGGAGTGATGGCAGCACGGGA 300
Db GCACAGACATTCCTGGAGCAGTGACCTTCTGCTGGAGGAGTGATGGCAGCACGGGA 335

QY 301 CAACTGGGACCCACTTGCCTCTCATCCCTCTGGGCGAGCTTCTGGACAGGTCCTGTC 360
Db CAACTGGGACCCACTTGCCTCTCATCCCTCTGGGCGAGCTTCTGGACAGGTCCTGTC 395

QY 361 CTCTTTGGGCGCTCGCAGAGCTCTTGGAAACCCAGCTTCTCCACAGGCGAGGACACA 420
Db CTCTTTGGGCGCTCGCAGAGCTCTTGGAAACCCAGCTTCTCCACAGGCGAGGACACA 455

QY 421 GCTCACAAGATCCCAATGCCATCTTCTGAGCTTCCACACCTGCTCCAGGAAAGGTG 480
Db GCTCACAAGATCCCAATGCCATCTTCTGAGCTTCCACACCTGCTCCAGGAAAGGTG 515

QY 481 CGTTTCTGATGCTGTAGGAGGTCACCTCTGCGTACGG 522
Db CGTTTCTGATGCTGTAGGAGGTCACCTCTGCGTACGG 557

RESULT 12
AR023468 AR023468 1342 bp DNA PAT 05-DEC-1998
LOCUS Sequence 24 from patent US 5795569.
DEFINITION AR023468
ACCESSION AR023468
VERSION AR023468.1 GI:3976762
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1342)
AUTHORS Bartley,T.D., Bogenberger,J.M., Bosselman,R.A., Hunt,P.,
Kinstler,O.B. and Samal,B.B.
TITLE Mono-pegylated proteins that stimulate megakaryocyte growth and
differentiation
JOURNAL Patent: US 5795569-A 24 18-AUG-1998;
FEATURES Location/Qualifiers
SOURCE 1..1342
/organism="unknown"
BASE COUNT 299 a 454 c 293 g 296 t
ORIGIN
Query Match 47.4%; Score 518.8; DB 6; Length 1342;
Best Local Similarity 99.6%; Pred. No. 5.5e-96;
Matches 520; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAGCTGACTGAATGCTCCGTCGTCATGCTTCTCCTAACTGAAGCTAACGCTG 60
Db ATGAGCTGACTGAATGCTCCGTCGTCATGCTTCTCCTAACTGAAGCTAACGCTG 95

QY 61 TCAGAGCCGGCTCCTCTGCTGTTGACCTCCGAGTCCAGTAACTGCTTCGTGACTCC 120
Db TCAGAGCCGGCTCCTCTGCTGTTGACCTCCGAGTCCAGTAACTGCTTCGTGACTCC 155

QY 121 CATGTCCTTCCACAGCAGACTGAGCCAGTGCAGAGTTCACCTTTGCTTACACCTGTC 180
Db CATGTCCTTCCACAGCAGACTGAGCCAGTGCAGAGTTCACCTTTGCTTACACCTGTC 215
```

```
|||||
Db 156 CATGTCCTTACAGCAGACTGAGCCAGTGCAGAGGTTTCAACCTTTGCTACACTGTC 215
Qy 181 CTGCTGCTCTGTGGACTTTAGCTTTGGGAGATGGAAACCCAGATGGAGGAGACCAAG 240
Db 216 CTGCTGCTCTGTGGACTTTAGCTTTGGGAGATGGAAACCCAGATGGAGGAGACCAAG 275
Qy 241 GCACAGACATTTCTGGAGAGTGAACCTTCTGCTGGAGGAGTGAATGCAGACACGGGA 300
Db 276 GCACAGACATTTCTGGAGAGTGAACCTTCTGCTGGAGGAGTGAATGCAGACACGGGA 335
Qy 301 CAACCTGGGACCCACTTGCCTCTCATCCCTCTGGGACAGCTTTCTGGACAGTCCGTC 360
Db 336 CAACCTGGGACCCACTTGCCTCTCATCCCTCTGGGACAGCTTTCTGGACAGTCCGTC 395
Qy 361 CTCTCTGGGACCCACTTGCCTCTCATCCCTCTGGGACAGCTTTCTGGACAGTCCGTC 420
Db 396 CTCTCTGGGACCCACTTGCCTCTCATCCCTCTGGGACAGCTTTCTGGACAGTCCGTC 455
Qy 421 GCTCACAAGGATCCCAATGCCATCTTCTGAGCTTCCAAACACCTGCTCCGAGGAAAGTG 480
Db 456 GCTCACAAGGATCCCAATGCCATCTTCTGAGCTTCCAAACACCTGCTCCGAGGAAAGTG 515
Qy 481 CGTTTCTGATCTTGTAGAGGTCACCCCTCTGGGTACGG 522
Db 516 CGTTTCTGATCTTGTAGAGGTCACCCCTCTGGGTACGG 557

RESULT 13
AR087573
LOCUS AR087573 1342 bp DNA PAT 07-SEP-2000
DEFINITION Sequence 1 from patent US 5989538.
ACCESSION AR087573
VERSION AR087573.1 GI:10014336
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1342)
AUTHORS Elliott,S.G.
TITLE Mpl ligand analogs
JOURNAL Patent: US 5989538-A 1 23-NOV-1999;
FEATURES Location/Qualifiers
source 1. 1342
BASE COUNT 299 a 454 c 293 g 296 t
ORIGIN

Query Match 47.4%; Score 518.8; DB 6; Length 1342;
Best Local Similarity 99.6%; Pred. No. 5.5e-96;
Matches 520; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGGAGCTGACTGAATGCTCTCTGCTGGTCAATGCTTCTTAACCTGCAAGGCTAACGCTG 60
Db 36 ATGGAGCTGACTGAATGCTCTCTGCTGGTCAATGCTTCTTAACCTGCAAGGCTAACGCTG 95
Qy 61 TCCAGCCCGGCTCTCTCTGCTGGTCAATGCTTCTTAACCTGCAAGGCTAACGCTG 120
Db 96 TCCAGCCCGGCTCTCTCTGCTGGTCAATGCTTCTTAACCTGCAAGGCTAACGCTG 155
Qy 121 CATGTCCTTACAGCAGACTGAGCCAGTGCAGAGGTTTCAACCTTTGCTACACCTGTC 180
Db 156 CATGTCCTTACAGCAGACTGAGCCAGTGCAGAGGTTTCAACCTTTGCTACACCTGTC 215
Qy 181 CTGCTGCTCTGTGGACTTTAGCTTTGGGAGATGGAAACCCAGATGGAGGAGACCAAG 240
Db 216 CTGCTGCTCTGTGGACTTTAGCTTTGGGAGATGGAAACCCAGATGGAGGAGACCAAG 275
Qy 241 TCCAGCCCGGCTCTCTCTGCTGGTCAATGCTTCTTAACCTGCAAGGCTAACGCTG 300
Db 276 TCCAGCCCGGCTCTCTCTGCTGGTCAATGCTTCTTAACCTGCAAGGCTAACGCTG 335
Qy 301 CATGTCCTTACAGCAGACTGAGCCAGTGCAGAGGTTTCAACCTTTGCTACACCTGTC 360
Db 336 CATGTCCTTACAGCAGACTGAGCCAGTGCAGAGGTTTCAACCTTTGCTACACCTGTC 395
Qy 361 CTGCTGCTCTGTGGACTTTAGCTTTGGGAGATGGAAACCCAGATGGAGGAGACCAAG 420
Db 396 CTGCTGCTCTGTGGACTTTAGCTTTGGGAGATGGAAACCCAGATGGAGGAGACCAAG 455
Qy 421 GCACAGACATTTCTGGAGAGTGAACCTTCTGCTGGAGGAGTGAATGCAGACACGGGA 300
Db 456 GCACAGACATTTCTGGAGAGTGAACCTTCTGCTGGAGGAGTGAATGCAGACACGGGA 335
```

```
Qy 301 CAACCTGGGACCCACTTGCCTCTCATCCCTCTGGGACAGCTTTCTGGACAGTCCGTC 360
Db 336 CAACCTGGGACCCACTTGCCTCTCATCCCTCTGGGACAGCTTTCTGGACAGTCCGTC 395
Qy 361 CTCTCTGGGACCCACTTGCCTCTCATCCCTCTGGGACAGCTTTCTGGACAGTCCGTC 420
Db 396 CTCTCTGGGACCCACTTGCCTCTCATCCCTCTGGGACAGCTTTCTGGACAGTCCGTC 455
Qy 421 GCTCACAAGGATCCCAATGCCATCTTCTGAGCTTCCAAACACCTGCTCCGAGGAAAGTG 480
Db 456 GCTCACAAGGATCCCAATGCCATCTTCTGAGCTTCCAAACACCTGCTCCGAGGAAAGTG 515
Qy 481 CGTTTCTGATCTTGTAGAGGTCACCCCTCTGGGTACGG 522
Db 516 CGTTTCTGATCTTGTAGAGGTCACCCCTCTGGGTACGG 557

RESULT 14
AR087573
LOCUS AR087573 1342 bp DNA PAT 04-APR-1998
DEFINITION Sequence 1 from patent US 5696250.
ACCESSION AR087573
VERSION AR087573.1 GI:3022649
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1342)
AUTHORS Elliott,S.G.
TITLE DNA encoding megakaryocyte growth and development factor analogs
JOURNAL Patent: US 5696250-A 1 09-DEC-1997;
FEATURES Location/Qualifiers
source 1. 1342
BASE COUNT 299 a 454 c 293 g 296 t
ORIGIN

Query Match 47.4%; Score 518.8; DB 6; Length 1342;
Best Local Similarity 99.6%; Pred. No. 5.5e-96;
Matches 520; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGGAGCTGACTGAATGCTCTCTGCTGGTCAATGCTTCTTAACCTGCAAGGCTAACGCTG 60
Db 36 ATGGAGCTGACTGAATGCTCTCTGCTGGTCAATGCTTCTTAACCTGCAAGGCTAACGCTG 95
Qy 61 TCCAGCCCGGCTCTCTCTGCTGGTCAATGCTTCTTAACCTGCAAGGCTAACGCTG 120
Db 96 TCCAGCCCGGCTCTCTCTGCTGGTCAATGCTTCTTAACCTGCAAGGCTAACGCTG 155
Qy 121 CATGTCCTTACAGCAGACTGAGCCAGTGCAGAGGTTTCAACCTTTGCTACACCTGTC 180
Db 156 CATGTCCTTACAGCAGACTGAGCCAGTGCAGAGGTTTCAACCTTTGCTACACCTGTC 215
Qy 181 CTGCTGCTCTGTGGACTTTAGCTTTGGGAGATGGAAACCCAGATGGAGGAGACCAAG 240
Db 216 CTGCTGCTCTGTGGACTTTAGCTTTGGGAGATGGAAACCCAGATGGAGGAGACCAAG 275
Qy 241 GCACAGACATTTCTGGAGAGTGAACCTTCTGCTGGAGGAGTGAATGCAGACACGGGA 300
Db 276 GCACAGACATTTCTGGAGAGTGAACCTTCTGCTGGAGGAGTGAATGCAGACACGGGA 335
Qy 301 CAACCTGGGACCCACTTGCCTCTCATCCCTCTGGGACAGCTTTCTGGACAGTCCGTC 360
Db 336 CAACCTGGGACCCACTTGCCTCTCATCCCTCTGGGACAGCTTTCTGGACAGTCCGTC 395
Qy 361 CTCTCTGGGACCCACTTGCCTCTCATCCCTCTGGGACAGCTTTCTGGACAGTCCGTC 420
Db 396 CTCTCTGGGACCCACTTGCCTCTCATCCCTCTGGGACAGCTTTCTGGACAGTCCGTC 455
Qy 421 GCTCACAAGGATCCCAATGCCATCTTCTGAGCTTCCAAACACCTGCTCCGAGGAAAGTG 480
Db 456 GCTCACAAGGATCCCAATGCCATCTTCTGAGCTTCCAAACACCTGCTCCGAGGAAAGTG 515
```

QY 481 CGTTTCCTGATGCTGTAGAGGGTCCACCTCTGCGTACGG 522
|||||
Db 516 CGTTTCCTGATGCTGTAGAGGGTCCACCTCTGCGTACGG 557
|||||

RESULT 15
D32047 1691 bp mRNA PRI 03-SEP-1997
LOCUS Homo sapiens mRNA for thrombopoietin, complete cds.
ACCESSION D32047
VERSION D32047.1 GI:2351117
KEYWORDS thrombopoietin.
SOURCE Homo sapiens liver cdna to mRNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Ogami, K.
Direct Submission
Submitted (08-JUL-1994) to the DDBJ/EMBL/GenBank databases. Kinya
Ogami, KIRIN Brewery Co., Ltd., Pharmaceutical Research Laboratory;
2-2 Soujamachi 1 chome, Maebashi, Gunma 371, Japan
(Tel:0272-54-8618, Fax:0272-52-2307)
2 (sites)
Kato, T., Ogami, K., Shimada, Y., Iwamatsu, A., Sohma, Y., Akahori, H.,
Horie, K., Kokubo, A., Kudo, Y., Maeda, E., Kobayashi, K., Ohashi, H.,
Ozawa, T., Inoue, H., Kawamura, K. and Miyazaki, H.
Purification and characterization of thrombopoietin
J. Biochem. 118 (1), 229-236 (1995)
96015174
Erratum: ([published erratum appears in J Biochem (Tokyo) 1996
Jan; 119(1):208])
Location/Qualifiers
1..1691
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="liver"
102..164
102..1163
/codon_start=1
/product="thrombopoietin"
/protein_id="BAA21930.1"
/db_xref="GI:2351118"
/translation="MELTELLVVMILLTARLTSSPPACDLRLVLSKLRDHSVH
SRLSQPEVPLEPVLPAVDFSLGKWKQMBETRAQDILGAVTLLEGVMAARGOL
GPTCLSLQLSGVRLGLGALQSLGTLQPPQGRTRAHKDPNAFLSFQHLRLGRV
RFLMLVGGSTLCLVRAPPTTAVFSRTSLVLTNLELNPRTSGLLFTASARTGSGL
LKWQQGFRAKIPGLNQTSRLDQIFGYLNRIHELLNGTRGLFPGPSRRTLGAPDISS
GTSDTGLPPLQPGYSPSPTHPTGQYTLPLPPTLPTPVVQLHPLLPDPSAPTPTP
TSPLLNTSYTHSONLSQEG"

BASE COUNT 383 a 542 c 371 g 395 t
ORIGIN

Query Match 47.4%; Score 518.8; DB 9; Length 1691;
Best Local Similarity 99.6%; Pred. No. 5.1e-96;
Matches 520; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGAGCTGACTGAATTGCTCCTCGTGTGATGCTTCCTCAACTGCAAGGCTAACGCTG 60
|||||
Db 102 ATGGAGCTGACTGAATTGCTCCTCGTGTGATGCTTCCTCAACTGCAAGGCTAACGCTG 161
|||||

QY 61 TCCAGCCGGCTCCTCGTGTGATGCTTCCTCAACTGCAAGGCTAACGCTG 120
|||||

Db 162 TCCAGCCGGCTCCTCGTGTGATGCTTCCTCAACTGCAAGGCTAACGCTG 221
|||||

QY 121 CATGCTCTTACAGCAGACTGAGCCAGTGCAGAGGTTACACCTTTGCCCTACACCTGTC 180
|||||

Db 222 CATGCTCTTACAGCAGACTGAGCCAGTGCAGAGGTTACACCTTTGCCCTACACCTGTC 281
|||||

QY 181 CTGCTGCTGCTGTGGACTTTTAGCTTGGAGATGGAACCCAGATGGAGAGACCAAG 240
|||||

Search completed: April 5, 2002, 10:29:40
Job time: 4773 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 5, 2002, 10:33:11 ; Search time 188.75 Seconds
(without alignments)
4973.622 Million cell updates/sec

Title: US-09-680-514-8

Perfect score: 1095

Sequence: 1 ATGGAGCTGACTGAATTGCT.....TACGCCACCTTCCCGAGCCC 1095

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_1101.*
1: /SIDS2/gcgdata/geneseq/geneseq/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1095	100.0	1095	17 AAT41788	Fusion peptide #3
2	1030.2	94.1	1083	17 AAT41787	Fusion peptide #2
3	987.4	90.2	1047	17 AAT41786	Fusion peptide #1
4	518.8	47.4	861	16 AAQ99552	Thrombopoietin cod
5	518.8	47.4	1062	16 AAT03941	Human thrombopoiet
6	518.8	47.4	1062	16 AAT04050	Sequence encoding
7	518.8	47.4	1062	17 AAT34852	Human thrombopoiet
8	518.8	47.4	1062	17 AAT37383	Human thrombopoiet
9	518.8	47.4	1062	17 AAT32591	Human thrombopoiet
10	518.8	47.4	1062	18 AAT85555	Human thrombopoiet
11	518.8	47.4	1062	19 AAV21696	Human thrombopoiet

12	518.8	47.4	1062	21 AAA51991	Human thrombopoiet
13	518.8	47.4	1086	17 AAT47958	Human thrombopoiet
14	518.8	47.4	1342	16 AAQ99704	Human MGDf-1 and M
15	518.8	47.4	1342	17 AAT36657	Native human mpl 1
16	518.8	47.4	1342	17 AAT10025	Human MGDf-1/2 cdn
17	518.8	47.4	1721	16 AAQ99554	Thrombopoietin cod
18	518.8	47.4	1721	21 AAA40191	Human wild type th
19	518.8	47.4	1774	21 AAA47793	Human thrombopoiet
20	518.8	47.4	1795	16 AAQ94107	hML cDNA. Homo sa
21	518.8	47.4	1795	18 AAT64318	Human thrombopoiet
22	517.2	47.2	600	17 AAT36658	Truncated human mp
23	517.2	47.2	605	20 AAX32813	Human truncated mp
24	517.2	47.2	1267	16 AAQ99553	Thrombopoietin cod
25	514	46.9	525	11 AAQ04482	Plasmid pASN6 enco
26	510.8	46.6	525	11 AAQ04484	Plasmid pASN145 en
27	506	46.2	525	11 AAQ04481	Plasmid pAS28 enco
28	495.8	45.3	1342	17 AAT33933	Human megakaryocyt
29	495.8	45.3	1342	19 AAV29068	Nucleotide sequenc
30	495.8	45.3	1342	20 AAX32812	Human mpl ligand e
31	490.8	44.8	1062	19 AAV55365	Flt3L/IgG2b/G-CSF
32	490.4	44.8	1044	19 AAV55367	Flt3L/IgG2b/G-CSF
33	489.2	44.7	1002	19 AAV55401	Flt3L/GS/G-CSF c10
34	489.2	44.7	1002	19 AAV55403	Flt3L/GS/G-CSF c10
35	489.2	44.7	1003	19 AAV70366	Human flt3 ligand
36	489.2	44.7	1003	19 AAV70393	Nucleic acid seque
37	489.2	44.7	1003	19 AAV55110	G-CSF-flt-3 ligand
38	489.2	44.7	1005	19 AAV70391	Nucleic acid seque
39	486.8	44.5	525	9 AAN80947	G-CSF gene isolate
40	486.8	44.5	1520	8 AAN70223	Plasmid pBRV2 inse
41	486.8	44.5	1520	8 AAN71320	Sequence encoding
42	486.8	44.5	1521	7 AAN60937	Plasmid pBRV2 inse
43	486.8	44.5	1525	9 AAN81478	Sequence encoding
44	486.8	44.5	1525	10 AAN91086	Plasmid pP12 contg
45	486	44.4	921	16 AAQ97174	PMON13056 DNA enco

ALIGNMENTS

RESULT 1
AAT41788
ID AAT41788 standard; DNA; 1095 BP.
XX
AC AAT41788;
XX
DT 01-JUL-1997 (first entry)
XX
DE Fusion peptide #3 having G-CSF and TPO activity.
XX
KW Fusion protein; human granulocyte colony stimulating factor; hg-CSF;
KW thrombopoietin; TPO; spacer peptide; blood platelet production;
KW leukocyte production; anaemia; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT sig_peptide 1..63
FT /*tag= a
FT mat_peptide 64..1095
FT /*tag= b
XX
PN WO9634016-A1.
XX
PD 31-OCT-1996.
XX
PF 26-APR-1996; 96WO-JP01157.
XX
PR 26-APR-1995; 95JP-0102625.
XX
PA (KYOW) KYOWA HAKKO KOGYO KK.
XX
PI Anazawa H, Konishi N, Shiotsu Y, Tamaoki T, Terasaki Y;
PI Uchida K, Yamasaki M, Yamashita K, Yokoi H;

QY	1	ATGGAGCTGACTGAATGCTTCTCTGCTGGTCATCGTCTCTCTACTGCAAGGCTTAACGCTG	60
Db	1	atggagctgactgaatgcttctctgctggctcatcgctctctctactaactgaaggctaaacgctg	60
QY	61	TCCAGCCCGGCTCCTCTGCTGTGTGACCTCCGAGTCTCTCAGTAAACTGCTTCTGTGACTCC	120
Db	61	tcagcccggtcctcctgctgtgacctcagtcctcagtaaaactgcttctgactcc	120
QY	121	CATGTCTTTCACAGCAGACTGAGCCAGTGCCTCCAGAGGTTACCTTTTGCTTACACCTGTTC	180
Db	121	catgtcttcaagcagactgagccagtgccagaggttcacaccttgcctacacctgtc	180
QY	181	CTGCTGCCCTGCTGGACTTACGTTGGGAGATGTAANAACCCAGATGGAGGAGACCAAG	240
Db	181	ctgctgccctgctggaacttagcttggagaaatgaaacccagatgaggagagacaaag	240
QY	241	GCACAGACATTTCTGGGAGCAGTGAACCTTCTCTGGAGGAGTGTATGGCAGCACGGGA	300
Db	241	gcacagacattctggagcagtgaccttctgtgagggagtgatggcagcaacgggga	300
QY	301	CAACTGGAGCCACTTGTCTCTCATCTCTCTGGGAGCGTTTCTTGACAGGTCGTCTCTC	360
Db	301	caactggagccacttgctctctcatctctctggcagcttctgtgagcaggttcgtctc	360
QY	361	CTCCTTTGGGCGCTGCAGAGCCCTCTTTGAAACCCAGCTTCTCTCCACAGGCGCAGACACA	420
Db	361	ctccttggggccctgcagagcctctcttggaaacccagcttctctccacagggcagacaca	420
QY	421	GCTCACAGGATCCCAATGCCATCTTCTCTGAGCTTCCACACCTGCTCCGAGGAAAGGTG	480
Db	421	gctcaagagatcccaatgccatctctctgagcttccaacacctgctccgaggaagggtg	480
QY	481	CGTTTCTCATGCTTTGAGGAGGTCACCCCTCTGCGTACGGTCCGGAGGTGCTCTGGC	540
Db	481	cgttctctgatgctgttagagggttccacctctgctcagg---ggtagcggttctgga	537
QY	541	GGTGTTCTGTGTGGCGGCTCCGGAGCGGTCTGTCCGCCAACATATCGCGCTCTCAGTCTA	600
Db	538	ggtgttctcgagggggttc-----tagagacccaacatatcgccctcagctcta	598
QY	601	CCACAGAGCTTCTTTTAAAGCTTAGAGCAAGTGAAGAGATCCAGGGCGATGCGCA	660
Db	589	ccacagagcttctctttaaaggcttagagcaagttaggaagatccaggcgcatggcgca	648
QY	661	GCGCTCCAGGAGAGCTGTGTGCCACTACAAGCTGTGCCACCCCGAGAGGCTGTGCTG	720
Db	649	gcgctccaaggagaagctgtgtgccactacaagctgtgccaccgcgaggagcttggcgctg	708
QY	721	CTGGACACTCTCTGGGCATCCCTCGGCTCCCTGAGCAGCTGCCCCAGCCAGGCGCCCTG	780
Db	709	ctcgacactctctgggcatacccttggctcccttgagcagctgtccccagcaggccctg	768
QY	781	CAGCTGGCAGGCTGCTTGAGCCACTCCATAGGGGCTTTTCTCTACACAGGGGCTCTTG	840
Db	769	cagctggcaggctgcttggaccaactccatagcgcttctctctaccaggggctcttg	828

Qy	841	CAGGCCCTGGAAGGATCTCCCCCGAGTTGGGTGCCACACTTGGACACACTTGGACCTGGAC	900
Db	829	caggccctggaaggatctcccccgagttgggtccccaccttggacacactgcagctggac	888
Qy	901	GTGCGCGACTTTCGCCACCCACCACTCTGCAGCAGCAGATGGAGAACTGGGAATGGCCCTGCCC	960
Db	889	gtcgcgagactttgccaccaccaactctggcgcgcagatggaaagaactgggaatggccccctgccc	948
Qy	961	CTGCAGGCCACCCAGGCGTGCATGCCGGCGCTTCGCCTCTGCTTCCAGCGCCGGGCGAGGA	1020
Db	949	ctgcagccacccaggggtgccatgcgcgccttcgcctctgcttcagcgcgcgggcagga	1008
Qy	1021	GGGCTCTAGTTGCCCTCCCATCTGCAGAGACTTCCTGGAGAGTCTGCTACCGGCTTCTACGC	1080
Db	1009	ggggctcagttgcttccccatctgcagagcttcctggagtgctgacccggttctacgc	1068
Qy	1081	CACCTTGGCCAGCCCC	1095
Db	1069	caacttggccccagccc	1083
RESULT	3		
AAAT41786			
ID	AAAT41786	standard; DNA; 1047 BP.	
XX	AAAT41786;		
XX			
DT	01-JUL-1997	(first entry)	
XX			
DE	Fusion peptide #1	having G-CSF and TPO activity.	
XX			
KW	Fusion protein; human granulocyte colony stimulating factor; hG-CSF;		
KW	thrombopoietin; TPO; spacer peptide; blood platelet production;		
KW	leukocyte production; anaemia; ds.		
XX			
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	sig_peptide	1..63	
FT		/*tag= a	
FT	mat_peptide	64..1047	
FT		/*tag= b	
XX			
PN	WO9634016-A1.		
PD	31-OCT-1996.		
XX			
PF	26-APR-1996;	96WO-JP01157.	
XX			
PR	26-APR-1995;	95JP-0102625.	
XX			
PA	(KYOW) KYOWA HAKKO KOGYO KK.		
XX			
PI	Anazawa H, Konishi N, Shiotsu Y, Tamaoki T, Terasaki Y;		
PI	Uchida K, Yamasaki M, Yamashita K, Yokoi H;		
XX			
DR	WPI: 1996-497573/49.		
DR	P-PSDB; AAW00377.		
XX			
PT	Fusion peptide having G-CSF and with thrombopoietin activity -		
PT	optionally chemically modified with a poly:alkylene glycol, used for		
PT	treatment of anaemia		
XX			
PS	Claim 3; Page 46-48; 71pp; Japanese.		
XX			
CC	The sequences given in AAT41786-88 encode fusion proteins which are		
CC	composed of a peptide having human granulocyte colony stimulating		
CC	factor (hG-CSF) activity fused with a peptide having thrombopoietin		
CC	(TPO) activity, opt. via a spacer peptide. Peptides derived from		
CC	these by deletion, insertion or substitution of one or more amino		
CC	acid residues are included within the scope of the invention. The		
CC	fusion peptides stimulate blood platelet and leukocyte production		
CC	and are useful in the treatment of anaemia.		

protein that stimulates megakaryopoiesis and thrombocytopoiesis. The cDNA can be incorporated into an expression vector and utilised in the prodn. of low mol.wt. bioactive human thrombopoietin in eukaryotic (partic. yeast or mammalian) host cells. The recombinant thrombopoietin has an N-terminus at Ser-22 and a C-terminus between Arg-185 and Asn-192 (esp. Arg-198 or Phe-207), and is useful for increasing platelet prodn. in the treatment of thrombocytopenia or for increasing proliferation of cells in bone marrow.

Sequence 1052 BP: 223 A; 375 C; 236 G; 228 T; 0 other:

```

RESULT 9
ID AAT32591 standard; cDNA; 1062 BP.
XX AC AAT32591;
XX 30-SEP-1996 (first entry)
XX DE Human thrombopoietin cDNA.
XX KW Thrombopoietin; TPO; tissue plasminogen activator; cytokine;
KW protein secretion; signal peptide; thrombocytopenia; therapy; ds.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
FT sig_peptide 1..63
FT /*tag= a
FT mat_peptide 64..1059
FT /*tag= b
XX PN WO9617067-A1.
XX 06-JUN-1996.
XX 15-NOV-1995; 95WO-US14932.
XX 30-NOV-1994; 94US-0347029.
XX (ZYMO ) ZYMOGENETICS INC.
XX PI Foster DC, Heipel MD, Holly RD;
XX WPI; 1996-277784/28.
XX P-PSDB; AAR99599.
XX DNA construct encoding native mammalian tissue plasminogen activator
XX secretory peptide and thrombopoietin polypeptide - used to produce
XX thrombopoietin polypeptide(s) on a large scale and in cost effective
XX manner
XX Disclosure; Page 41-43; 56pp; English.
XX A cDNA clone (AAT32591) codes for human thrombopoietin (TPO) (AAR99599).
XX Novel DNA constructs code for a secretory peptide (AAR9600) fused to
XX the TPO polypeptide, with a cleavage site at the junction. Such
XX constructs can be expressed in transformed host cells, pref.
XX Pichia pastoris, Saccharomyces cerevisiae or rodent kidney cells,
XX for large-scale, cost-effective prodn. of human TPO, useful for the
XX treatment of thrombocytopenia.
XX Sequence 1062 BP; 223 A; 375 C; 236 G; 228 T; 0 other;

Query Match 47.4%; Score 518.8; DB 17; Length 1062;
Best Local Similarity 99.6%; Pred. No. 1.4e-112; Indels 0; Gaps 0;
Matches 520; Conservative 0; Mismatches 2;

QY 1 ATGAGCTGACTGAATGTCTCGTGGTCATGCTTCTCTTAACAGCAAGGCTAACGCTG 60
DB 1 atggagctgactgaattgtctctcggtgcatgcttctcttaactgaagcgaagcgtg 60
QY 61 TCAGAGCCGCTCTCTGCTGTGTGACATCGAGTCCTCAGTAACCTGCTTGCTGACTCC 120
DB 61 tcagagccgctctctgctgtgtgacatcgagtcctcagtaaacctgcttgctgactcc 120
QY 121 CATGTCTCTTCACAGCAGACTGAGCCAGTCCCGCAGAGGTTTCACCTTTGCTACACTGTC 180
DB 121 catgtctcttcacagcagactgagccagtgcccgagaggttcacaccttgcctcacactgctc 180
QY 181 CTGCTGCTGCTGTGGACTTTAGCTTTGGGAGAAATGGAAACCCAGATGGAGGAGACCAAG 240
DB 181 ctgctgctgctgtggaacttttagctttgggagaaatggaaacccagatggaggaagaccag 240
QY 241 GCACAGGACATTTCTGGGAGCAGTGACCTTCTGCTGAGGAGGAGTGTGGCAGCAGCGGGA 300
DB 241 gcaagagacattctggagcagtgacccttctgctgagggagtgatggcagcacgggga 300
QY 301 CAACTGGGACCCACTTGCCTCTCATCCCTCCCTGGGCGAGCTTTTCTGGACAGGTCCTGCTC 360
DB 301 caactgggacccacttgcctctcatccctcctctgggagcagcttcttgagacaggtccgtctc 360
QY 361 CTCTTTGGGGCCCTGCAGAGCCTCTCTTGAACCCAGCTTCTCTCCACAGGCGAGACCAACA 420
DB 361 ctctttggggccctgcagagcctcttggaaacccagcttctctccacagggcagagaccaca 420
QY 421 GCTCACAGGATCCCAATGCCATCTTCTCTGAGCTTCCACACACCTGCTCCGAGGAAAGTG 480
DB 421 gctcaagagatcccaatgccatcttctctgagcttccacacacctgctccgaggaaggtg 480
QY 481 CGTTTCTCTGATGCTTGTAGGAGGTCACCTCTCTGCTGACGG 522
DB 481 cgttctctgatgcttgttaggaggggtccacctctgctgctcagg 522

RESULT 10
AAT85555
ID AAT85555 standard; cDNA; 1062 BP.
XX AC AAT85555;
XX 05-NOV-1997 (first entry)
XX DE Human thrombopoietin cDNA.
XX KW TPO; cancer; myeloid cell proliferation; expansion; bone marrow;
KW peripheral blood stem cell; chemotherapy; radiation therapy;
KW breast cancer; leukaemia; lymphoma; multiple myeloma; ds.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
FT CDS 1..1062
FT /*tag= a
FT mat_peptide 64..1059
FT /*tag= b
FT /*product= Thrombopoietin
XX WO9640218-A1.
XX 19-DEC-1996.
XX 29-MAY-1996; 96WO-US07880.
XX 07-JUN-1995; 95US-0482212.
XX (ZYMO ) ZYMOGENETICS INC.
XX PI Fibbe WE, Grossmann A;
XX WPI; 1997-108626/10.
XX P-PSDB; AAW26841.
XX Increasing haematopoietic cells in patient - by admin. of stem cells
XX from donor previously treated with thrombopoietin, used in patients
XX who have received chemotherapy or radiation for cancer
XX Disclosure; Page 17-19; 32pp; English.
XX In a claimed method, thrombopoietin (TPO) (especially human TPO) is
XX administered to a donor to stimulate proliferation of myeloid cells.
XX Bone marrow or peripheral blood stem cells are collected from the
XX donor and administered to the recipient. The method is used for
XX preparing cells for (bone marrow) transplantation and for stimulating

```

CC platelet or erythrocyte recovery in a patient receiving chemotherapy
CC or radiation therapy for e.g. breast cancer, leukaemia, lymphoma or
CC multiple myeloma. Treating the donor with TPO accelerates restoration
CC of erythrocyte and thrombocyte levels in the patient after transplant.
CC This restoration may be further improved by administering TPO to the
CC recipient after transplant. The present sequence encodes human TPO.
XX
XX
SQ Sequence 1062 BP; 223 A; 375 C; 236 G; 228 T; 0 other;

Query Match 47.4%; Score 518.8; DB 18; Length 1062;
Best Local Similarity 99.6%; Pred. No. 1.4e-112;
Matches 520; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAGCTGACTGAATGCTCTCGTGGTCATGCTCTCTAACTCAAGGCTAAGCGTG 60
DB 1 atggagctgactgaatgctctcg99tcatgctctctcaactgcaaggctaaagcgtg 60
QY 61 TCCAGCCCGGCTCCTCCTGCTGTGACCTCCGAGTCTCAGTAAACTGCTTCGTGACTCC 120
DB 61 tccagcccggtcctcctgctgtgacctccgagtcctcagtaaaactgctctgactcc 120
QY 121 CATGCTCTCACAGCAGACTGACCCAGTGCAGAGGTTCCACCTTTGCCCTACACCTGTC 180
DB 121 catgctctcacagcagactgagccagtgcccagaggttcacaccttgctacacctgtc 180
QY 181 CTGCTGCTCTGCTGACTTTAGCTTGGGAGATGGAACCCAGATGGAGGAGACCAAG 240
DB 181 ctgctgctctgctgactttagcttgggagatggaacccagatggagggagaccaag 240
QY 241 GCACAGGACATTCCTGGAGCAGTGACCCCTTCTGCTGGAGGAGTGATGGCAGCACGGGA 300
DB 241 gcacaggacattctggagcagtgaccttctgctgagggagtatggcagcacggga 300
QY 301 CAACTGGGACCACTTGCTCTCTATCCCTGCTGGGAGCTTCTTGACAGGTCGCTCTC 360
DB 301 caactgggaccacttgctctctatccctcctctgggagcttcttgagcaggtccgtctc 360
QY 361 CTCTTGGGGCCCTGAGAGCCCTTGGAAACCCAGCTTCTCCACAGGCGAGGACACA 420
DB 361 ctcttggggccctgagagcccttggaaacccagcttctccacagggagaccaca 420
QY 421 GCTCAAGGATCCCAATGCTCTCTGAGCTTCCACACCTGCTCCGAGGAAAGGTG 480
DB 421 gctcaaggatcccaatgctctctgagcttccacacctgctccagaggaagggtg 480
QY 481 CGTTTCTGATGCTTGTAGAGGCTCCACCTCTGCGTAGCG 522
DB 481 cgttctctgctgttaggaggtccacctctgctcagg 522

RESULT 11
AAV21696
ID AAV21696 standard; cDNA; 1062 BP.
XX
AC AAV21696;
XX
DT 07-AUG-1998 (first entry)
XX
DE Human thrombopoietin encoding cDNA.
XX
KW Thrombopoietin; TPO; expression vector; platelet number; cytokine;
KW human; bone marrow; proliferation; treatment; cancer; aplastic anaemia;
KW myelodysplastic syndrome; chemotherapy; cytopenia; thrombocytopenia;
KW haematologic disorder; leukaemia; lymphoma; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..1062
FT /tag= a
FT /product= "Thrombopoietin polypeptide"
FT sig_peptide 1..63

FT mat_peptide /tag= b
FT 64..1059
FT /tag= c
XX
PN W09806849-A1.
XX
PD 19-FEB-1998.
XX
PF 30-JUL-1997; 97WO-US13543.
XX
PR 13-AUG-1996; 96US-0696447.
XX
PA (ZYMO) ZYMOGENETICS INC.
XX
PI Irani M, Morrison-nelson GR;
XX
XX WPI; 1998-159541/14.
DR P-PSDB; AAW53124.
XX
XX Thrombopoietin protein expression vector - used for increasing
PT platelet number in a mammal
PS Disclosure; Pages 29-32; 56pp; English.
XX
CC This cDNA encodes a thrombopoietin (TPO) polypeptide. This can be used
CC in the construction of a new expression vector replicable in a
CC eukaryotic host cell encoding TPO polypeptides. The vector comprises a
CC transcription promoter, a first DNA segment encoding a secretory leader,
CC a second segment encoding a TPO polypeptide and a transcription
CC terminator which are all operably linked. The second TPO segment consists
CC of C-X-B, where C is a human TPO cytokine domain, X is a peptide bond or
CC a linker consisting of one or two amino acid residues, such that X along
CC in combination with C or B does not provide a dibasic amino acid pair
CC and B is a polypeptide that can be selected from AAW53125 to AAW53126.
CC The secretory leader is a S. cerevisiae alpha-factor secretory leader. A
CC cultured yeast cell containing such an expression vector can be used to
CC produce the TPO polypeptide. The TPO polypeptide can be used in a method
CC for increasing platelet number in a mammal. It can be used to increase
CC proliferation of bone marrow cells for treatment of cytopenia, including
CC those induced by aplastic anaemia, myelodysplastic syndromes, including
CC chemotherapy or congenital cytopenias. It can also be used to treat
CC thrombocytopenia, haematologic disorders, such as leukaemia and lymphoma
CC or metastatic cancers involving bone marrow.
XX
SQ Sequence 1062 BP; 223 A; 375 C; 236 G; 228 T; 0 other;

Query Match 47.4%; Score 518.8; DB 19; Length 1062;
Best Local Similarity 99.6%; Pred. No. 1.4e-112;
Matches 520; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAGCTGACTGAATGCTCTCGTGGTCATGCTCTCTAACTCAAGGCTAAGCGTG 60
DB 1 atggagctgactgaatgctctcg99tcatgctctctcaactgcaaggctaaagcgtg 60
QY 61 TCCAGCCCGGCTCCTCCTGCTGTGACCTCCGAGTCTCAGTAAACTGCTTCGTGACTCC 120
DB 61 tccagcccggtcctcctgctgtgacctccgagtcctcagtaaaactgctctgactcc 120
QY 121 CATGCTCTCACAGCAGACTGACCCAGTGCAGAGGTTCCACCTTTGCCCTACACCTGTC 180
DB 121 catgctctcacagcagactgagccagtgcccagaggttcacaccttgctacacctgtc 180
QY 181 CTGCTGCTCTGCTGACTTTAGCTTGGGAGATGGAACCCAGATGGAGGAGACCAAG 240
DB 181 ctgctgctctgctgactttagcttgggagatggaacccagatggagggagaccaag 240
QY 241 GCACAGGACATTCCTGGAGCAGTGACCCCTTCTGCTGGAGGAGTGATGGCAGCACGGGA 300
DB 241 gcacaggacattctggagcagtgaccttctgctgagggagtatggcagcacggga 300
QY 301 CAACTGGGACCACTTTGCTCTCTATCCCTGCTGGGAGCTTCTTGACAGGTCGCTCTC 360
DB 301 caactgggaccacttgctctctatccctcctctgggagcttcttgagcaggtccgtctc 360

PN W09628181-A1.
XX 19-SEP-1996.
XX 14-MAR-1996; 96WO-JP00635.
XX 15-MAR-1995; 95JP-0056248.
XX (KIRI) KIRIN BREWERY KK.
XX Otsuki N;
XX WPI; 1996-433541/43.
XX P-PSDB; AAW09314.
XX Stable thrombopoietin compositions contains TPO and an additive -
XX does not lose thrombopoietin activity on storage in a container
XX
XX Example 1; Page 26-28; 38pp; Japanese.
XX
XX Stable thrombopoietin (TPO) compsns. containing TPO and an additive,
XX pref. human serum albumin and/or gelatine, etc., are claimed. The
XX compsns. do not lose TPO activity on storage in a container as the
XX compsns. has reduced adsorption onto the container walls. The present
XX sequence encodes recombinant immature human TPO (1-332).
XX
XX Sequence 1086 BP; 229 A; 386 C; 243 G; 228 T; 0 other;

Query Match 47.4%; Score 518.8; DB 17; Length 1086;
Best Local Similarity 99.6%; Pred. No. 1.4e-112;
Matches 520; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGAGCTGACTGAATGCTCCGTCGTGTCATGCTTCTTAAGTCAAGGCTAACCGTG 60
DB 25 atggagctgactgaatgctccgtggtgatgcttcttaagtcgaaggtcaacgctg 84
QY 61 TCCAGCCCGGCTCCTCCTGCTGTGTGACCTCCGAGTCCCTAGTAACTGCTTGTGACTCC 120
DB 85 tccagcccgctcctcctggtgacctccgagtcctcagtaaacgttctgactcc 144
QY 121 CATGCTCTTACAGCAGCTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
DB 145 catgctcttcaagcagactgagcagtgccagaggttcaccccttgcctacacctgtc 204
QY 181 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
DB 205 ctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 264
QY 241 GCACAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
DB 265 gcacagcagctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 324
QY 301 CACTGGGACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
DB 325 caactgggacccactgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 384
QY 361 CTCTTGGGGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
DB 385 ctcttggggccctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 444
QY 421 GCTCACAGGATCCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
DB 445 gctcacaggatcccatgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 504
QY 481 CGTTTCTGATGCTTGTAGAGGGTCCACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 522
DB 505 cgtttctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 546

RESULT 14
AAQ99704
ID AAQ99704 standard; cDNA; 1342 BP.

XX AAQ99704;
XX 01-FEB-1996 (first entry)
XX Human MGDF-1 and MGDF-2.
XX Mpl ligand; mammalian megakaryocyte growth promoting factor;
XX platelet producing factor; ss.
XX Homo sapiens.
XX Key Location/Qualifiers
XX CDS 36..1097
XX FT /*tag= a
XX EP675201-A.
XX 04-OCT-1995.
XX 30-MAR-1995; 95EP-0104711.
XX 30-NOV-1994; 94US-0347780.
XX 31-MAR-1994; 94US-0221768.
XX 12-OCT-1994; 94US-0252628.
XX 12-OCT-1994; 94US-0321488.
XX (AMGE-) AMGEN INC.
XX Bartley TD, Bogenberger JM, Bosselman RA, Hunt P;
XX Kinstler OB, Samal BB;
XX WPI: 1995-338287/44.
XX P-PSDB; AAR80821, AAR80822, AAR80823, AAR80824, AAR80825, AAR80826.
XX New human megakaryocyte growth and development factor - used to increase
XX the number of megakaryocytes or platelets in patients, e.g. for treating
XX thrombocytopenia
XX Claim 17; Fig 11; 98pp; English.
XX Mpl ligand is synonymous with MGDF. Mpl ligands are mammalian
XX megakaryocyte growth promoting and/or platelet producing factors.
XX MGDF-2 is a truncated portion of MGDF-1. It is hypothesized that
XX human MGDF is expressed in vivo as a substantially inactive or less
XX active precursor polypeptide that contains variable C-terminal AAs.
XX upon cleavage of the C-terminal AAs (as well as the signal peptide),
XX the processed form(s) of the molecule retain activity or become more
XX active. It is believed that MGDF-1 may require processing in order to
XX exhibit its activity. The fact that a truncated form MGDF-1 (i.e.
XX MGDF-2) is active supports this hypothesis. Various active molecules
XX that may result from truncations of the sequence set forth as
XX MGDF-1 (AAR80824) are given in AAR80822, AAR80823, & AAR80825. AAR80826
XX is the signal peptide. The preferred truncation variants of MGDF-1
XX are any of those that have C-terminal truncations from AA 173-353
XX (along with cleavage of the signal peptide). The signal peptide
XX may have 23 AAs.
XX Sequence 1342 BP; 299 A; 454 C; 293 G; 296 T; 0 other;

Query Match 47.4%; Score 518.8; DB 16; Length 1342;
Best Local Similarity 99.6%; Pred. No. 1.5e-112;
Matches 520; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGAGCTGACTGAATGCTCCGTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 60
DB 36 atggagctgactgaatgctccgtgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 95
QY 61 TCCAGCCCGGCTCCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
DB 96 tccagcccgctcctcctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 155

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 5, 2002, 09:47:02 ; Search time 1909.33 Seconds

(without alignments)
6162.700 Million cell updates/sec

Title: US-09-680-514-8

Perfect score: 1095

Sequence: 1 ATGGAGCTGACTGAATGCT.....TAGGCACCTGCCAGGCC 1095

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: em_estfun:*
2: em_esthum:*
3: em_estin:*
4: em_estom:*
5: em_estpl:*
6: em_estba:*
7: em_estro:*
8: em_estov:*
9: em_htc:*
10: gb_estl:*
11: gb_est2:*
12: gb_htc:*
13: gb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_rod:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	320.4	29.3	973	11	BI411128
2	312.8	28.6	598	11	BG548320
3	301.6	27.5	900	11	BI328788
4	300.6	27.5	507	10	BE485194
5	300.6	27.5	633	10	AI314551
6	299.6	27.4	609	10	AA538257
7	285.2	26.0	878	11	BF789308
8	227.8	20.8	464	10	BE754587
9	135.2	12.4	327	11	BF848766
10	135.2	12.3	449	13	AZ025993
11	106.2	9.7	419	13	AZ352219
12	86.4	7.9	708	13	AZ792988

13	77.8	7.1	752	13	AQ888740
14	65.6	6.0	589	10	BE377769
15	54.8	5.0	925	13	CNS0091P
16	52.4	4.8	976	11	BG823538
17	50.2	4.6	513	10	BE480590
18	49	4.5	925	13	CNS0091P
19	48.8	4.5	617	10	AI981598
20	44.6	4.1	702	11	BG433224
21	43.6	4.0	296	10	AI556398
22	43.2	3.9	461	10	BE598945
23	43.2	3.9	559	11	BG556176
24	42.8	3.9	427	11	BG704075
25	42.6	3.9	452	10	AA155632
26	42.6	3.9	899	13	CNS01VVT
27	42.4	3.9	1203	13	CNS015V4
28	42.2	3.9	598	10	AA115932
29	42	3.8	507	10	AI712717
30	42	3.8	553	10	AW253985
31	41.4	3.8	421	10	AA058743
32	41.4	3.8	478	11	BI345642
33	40.8	3.7	545	11	BF868325
34	40.6	3.7	447	11	BG604563
35	40.6	3.7	845	10	AL572931
36	40.6	3.7	932	13	CNS0072Q
37	40.6	3.7	1003	11	BI410408
38	40.6	3.7	1101	13	CNS05709
39	40.4	3.7	445	10	AW732737
40	40.4	3.7	458	10	AW273202
41	40.4	3.7	485	10	BE463718
42	40.4	3.7	490	10	AW410572
43	40.4	3.7	593	11	BG519711
44	40.2	3.7	864	13	CNS022LE
45	39.8	3.6	387	10	AW193322

ALIGNMENTS

RESULT. 1	BI411128	LOCUS	602962472F1	NCI_CGAP_Lu33	Mus musculus	cdna	clone	IMAGE:5118185	5',
BI411128	EST	14-AUG-2001							
DEFINITION	602962472F1	NCI_CGAP_Lu33	Mus musculus	cdna	clone	IMAGE:5118185	5',		
ACCESSION	BI411128	GI:15172051							
VERSION	BI411128.1	GI:15172051							
KEYWORDS	EST.								
SOURCE	house mouse.								
ORGANISM	Mus musculus								
REFERENCE	1	(bases 1 to 973)							
AUTHORS	NIH-MGC	http://mgc.nci.nih.gov/							
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)								
JOURNAL	Unpublished (1999)								
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgabs-r@mail.nih.gov Tissue Procurement: Gilbert Smith, Ph.D. CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov Plate: LLAM11287 row: O column: 18 High quality sequence start: 32 High quality sequence stop: 912. Location/Qualifiers 1. .973 /organism="Mus musculus" /strain="CZECH II" /db_xref="taxon:10090" /clone="IMAGE:5118185"								
FEATURES	source								

Email: tadsepsi@parc.usda.gov
Single pass sequencing. Bases called and alt_trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 18 and -minmatch 12 options.

Query Match 27.4%; Score 299 6; DB 10; Length 609;
Best Local Similarity 81.6%; Pred. No. 4.7e-57;
Matches 364; Conservative 0; Mismatches 69; Indels 13; Gaps 1;

Qy	1	ATGAGCTGACTGAATTCCTCCTCGTGGTCATGCTTCTCCTAACTGCGCAAGCTAACGCTG	60
Db	177	ATGAGCTGACTGAATTCCTCCTCGGCGCCATGCTTCTGCAAGTGGCAAGACTAACTCTG	236
Qy	61	TCCAGCCCGGCTCCTCCTGCTTGTGACCTCCGAGTCCCTCAGTAAACTGCTTCGTGACTCC	120
Db	237	TCCAGCCCGTAGCTCCTGCGCTGTACCCAGACTCCTAAATAAATGCTCGCTGACTCC	296
Qy	121	CATGCTCTTACAGCAGACTCAGCCAGTGCAGCCAGAGTTTACCCCTTTTGCTTACACCTGTC	180
Db	297	CACCTCTCTTACAGCCGCACTGAGTCAGTGTCTCCGACGTCGACCCCTTTGTCTATCCCCGT	356
Qy	181	CTGCTGCTGCTGTGGACTTTAGCTTGGGAGAATGGAAAACCCAGATGGAGGAGACCAAG	240
Db	357	CTGCTGCTGCTGTGGACTTTAGCTTGGGAGAATGGAAAACCCAGACGAACAGACAAG	416
Qy	241	GCACAGGACATTCCTGGAGCAGTCAACCCCTTCTGCTGGAGGGAGTGTGGCAGCAGGGGA	300
Db	417	GCACAGGACATTCCTAGGGGCACTGCTCCTTCTACTGGAGGGAGTGTGGCAGCAGGAGA	476
Qy	301	CAACTGGACCCACTTGCCTCTCATCTCCTCTGGGGCAGCTTCTTGACAGGTCCTGCTC	360
Db	477	CAGTGGACCCCTCTGCTCTCATCTCCTCTGGGACAGCTTCTTGCGCAGGTCGCGCTC	536
Qy	361	CTCCTTGGGGCCCTGCAGAGCCTCCTTGGAAACCCAGCTTCTCCACAGGGCAGGACACA	420
Db	537	CTCTTGGGGGCCCTGCACCGCTCTTAAGGAA-----CCAGGGCAGGACCAACA	583
Qy	421	GCTCACAAGGATCCCAATGCCATCTT	446
Db	584	GCTCACAAGGACCCCAATGCCCTCTT	609

```

DEFINITION      602104729F1 NCI_CGAP_Kid14 Mus musculus cDNA clone IMAGE:422927
                  5', mRNA sequence.
ACCESSION      BF789308
VERSION        BF789308.1 GI:12094344
KEYWORDS       EST.
SOURCE         house mouse.
ORGANISM       Mus musculus
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE      NIH-MGC http://mgc.nci.nih.gov/.
                1 (bases 1 to 878)
                National Institutes of Health, Mammalian Gene Collection (MGC)
                Unpublished (1999)
COMMENT        Contact: Robert Strausberg, Ph.D.
                Email: cgapps-femail.nih.gov
                Tissue Procurement: Jeffrey E. Green, M.D.
                cDNA Library Preparation: Life Technologies, Inc.
                cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
                DNA Sequencing by: Incyte Genomics, Inc.
                Clone distribution: MGC clone distribution information can be
                found through the I.M.A.G.E. Consortium/LLNL at:
                http://image.llnl.gov
                Plate: LLAM9810 row: 1 column: 08
                High quality sequence source: 628.
FEATURES       1. .878
                Location/Qualifiers
                /organism="Mus musculus"
                /strain="FVB/N"
                /db_xref="taxon:10090"
                /clone="IMAGE:422927"
                /clone_lib="NCI_CGAP_Kid14"
source

```

[illegible]

Email: szhao@tigr.org
Clones are derived from the mouse BAC library RPCI-23. For BAC

library availability, please contact Pieter de Jong
(pieter@jeong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (<http://bacpac.med.buffalo.edu/orderingframe.htm>)
or from Resea ch Genetics (info@resgen.com). BAC end page:
http://www.tigr.org/cdb/bac_ends/mouse/bac_end_intro.html
Plate: 343 row: I column: 16
Seq primer: SP6
Class: BAC ends.

FEATURES

Location/Qualifiers
1. .449
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-23-343I16"
/clone_lib="RPCI-23"
/sex="Female"
/lab_host="DH10B"

/note="Organ: Kidney/Brain; Vector: pBACE3.6; Site:1:
EcoRI; Site:2: EcoRI; Female C57BL/6J mouse kidney and/or
brain genomic DNA was isolated and partially digested
with a combination of EcoRI and EcoRI Methylase. Size
selected DNA was cloned into the pBACE3.6 vector at the
EcoRI sites. The ligation products were transformed into
DH10B electrocompetent cells (BRL Life Technologies)."
90 a 121 c 108 g 129 t 1 others

BASE COUNT

Query Match 12.3%; Score 135.2; DB 13; Length 449;
Best Local Similarity 84.4%; Pred. No. 2.7e-20;
Matches 152; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY 228 GGAGGAGCAGGACAGACATCTGGAGCAGTGACCTTCTCTGGAGGAGTGAT 287
Db 126 GGAACAGAGCAGGACAGACATCTAGGGCAGTGCTTTTCTACTGGAGGAGTGAT 185

QY 288 GCGAGCAGGGGACACTGGAGCCACTTGCCTCTCATCCTCTCTGGGCGAGCTTCTGG 347
Db 186 GCGAGCAGGAGCAGTGGAACCTTCTGCTCTATCCTCTCTGGGAGCTTCTGG 245

QY 348 ACAGGTGCGTCTCTCTTGGGGCCCTGAGAGCCCTCTTGGAAACCCAGCTTCTCCACA 407
Db 246 GCAGGTTCGCTCTCTTGGGGCCCTGAGAGCCCTCTAGGAACCCAGTAAAGTCCCA 305

RESULT 11

AZ352219/c
LOCUS AZ352219 419 bp DNA GSS 29-SEP-2000
DEFINITION 1M0090P05R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC1M0090P05 R, DNA sequence.

ACCESSION AZ352219
VERSION AZ352219.1 GI:10431456
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 419)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.

REFERENCE 1 (bases 1 to 419)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.

REFERENCE 1 (bases 1 to 419)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.

REFERENCE 1 (bases 1 to 419)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.

REFERENCE 1 (bases 1 to 419)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.

REFERENCE 1 (bases 1 to 419)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.

REFERENCE 1 (bases 1 to 419)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.

Insert Length: 10000 Std Error: 0.00
Plate: 0090 row: P column: 05
Seq primer: CACACAGGAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 419.
Location/Qualifiers
1. .419

FEATURES

Location/Qualifiers
1. .419
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0090P05"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"

/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(<http://www.jax.org/resources/documents/dnares/>). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pWD42 (gil4732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid RL. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT 100 a 98 c 124 g 97 t
ORIGIN

Query Match 9.7%; Score 106.2; DB 13; Length 419;
Best Local Similarity 79.2%; Pred. No. 8.1e-14;
Matches 126; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 360 CCTCTCTGGGGCCCTCAGAGCCCTCTTGGAAACCCAGCTTCTCCACAGGCGAGGCCAC 419
Db 292 CCTACTCACTGGCCTCAGGCCCATCTCTGCCCTCAGCTTCTCTACAGGGCAGGCCAC 233

QY 420 AGCTCAAGAGTCCCAATGCCATCTTCTGAGCTTCCACACCTCTCCGAGGAAGGT 479
Db 232 AGCTCAAGAGTCCCAATGCCATCTTCTGAGCTTCCACACCTCTCCGAGGAAGGT 173

QY 480 GCGTTTCTGTGCTGTGTAGGAGGGTCCACCTCTCGCT 518
Db 172 GCGTTTCTGTGCTGTGTAGGAGGGTCCACCTCTCGCT 134

RESULT 12

AZ792988
LOCUS AZ792988 708 bp DNA GSS 16-FEB-2001
DEFINITION 2M0046105F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC2M0046105 F, DNA sequence.

ACCESSION AZ792988
VERSION AZ792988.1 GI:12937676
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 708)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.

REFERENCE 1 (bases 1 to 708)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.

REFERENCE 1 (bases 1 to 708)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.

REFERENCE 1 (bases 1 to 708)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.

REFERENCE 1 (bases 1 to 708)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.
